



Article A Frequency and Voltage Coordinated Control Strategy of Island Microgrid including Electric Vehicles

Peixiao Fan¹, Song Ke^{1,*}, Salah Kamel², Jun Yang¹, Yonghui Li¹, Jinxing Xiao³, Bingyan Xu³ and Ghamgeen Izat Rashed¹

- ¹ School of Electrical Engineering and Automation, Wuhan University, Wuhan 430072, China; whufpx0408@163.com (P.F.); JYang@whu.edu.cn (J.Y.); whusee2006@yahoo.com (Y.L.); ghamgeen@whu.edu.cn (G.I.R.)
- ² Electrical Engineering Department, Faculty of Engineering, Aswan University, Aswan 81542, Egypt; skamel@aswu.edu.eg
- ³ Electric Power Company of Shanghai, State Grid, Shanghai 200122, China; xjx1122@163.com (J.X.); gh197493@yahoo.com (B.X.)
- * Correspondence: kesong1997@whu.edu.cn; Tel.: +18-771-043566

Abstract: Frequency and voltage deviation are important standards for measuring energy indicators. It is important for microgrids to maintain the stability of voltage and frequency (VF). Aiming at the VF regulation of microgrid caused by wind disturbance and load fluctuation, a comprehensive VF control strategy for an islanded microgrid with electric vehicles (EVs) based on Deep Deterministic Policy Gradient (DDPG) is proposed in this paper. First of all, the *SOC* constraints of EVs are added to construct a cluster-EV charging model, by considering the randomness of users' travel demand and charging behavior. In addition, a four-quadrant two-way charger capacity model is introduced to build a microgrid VF control model including load, micro gas turbine (MT), EVs, and their random power increment constraints. Secondly, according to the two control goals of microgrid frequency and voltage, the structure of DDPG controller is designed. Then, the definition of space, the design of global and local reward functions, and the selection of optimal hyperparameters are completed. Finally, different scenarios are set up in an islanded microgrid with EVs, and the simulation results are compared with traditional PI control and $R(\lambda)$ control. The simulation results show that the proposed DDPG controller can quickly and efficiently suppress the VF fluctuations caused by wind disturbance and load fluctuations at the same time.

Keywords: islanded microgrid; electric vehicles; charger capacity model; VF control; DDPG

1. Introduction

Microgrid refers to a small power generation and distribution system that is composed of distributed power sources, energy storage devices, energy conversion devices, related loads, monitoring, and protection devices. It is an autonomous system that can realize self-control, protection, and management. In addition, the microgrid can operate in gridconnected mode and islanded mode. In islanded mode, the power quality of the microgrid is usually maintained by the micro sources and flexible loads [1]. At the same time, with the development of vehicle-to-grid (V2G) technology, the research of EVs in the areas of grid peak and valley filling, suppression of power fluctuations, and microgrid stability control has also been deepened [2,3], which brings opportunities and challenges to the VF regulation of microgrids.

Due to the limited capacity of the islanded microgrid, ensuring the stability of frequency and voltage is the key for the operation safety of microgrid. In [4], a VF strategy of an islanded microgrid based on fuzzy logic controller is proposed, which can control active and reactive powers and decrease power losses of the microgrid, thus the effectiveness and robustness of the proposed controller over the conventional proportional- integral controller. In [5], a decoupled VF controller for DGs is proposed, which is able to keep the



Citation: Fan, P.; Ke, S.; Kamel, S.; Yang, J.; Li, Y.; Xiao, J.; Xu, B.; Rashed, G.I. A Frequency and Voltage Coordinated Control Strategy of Island Microgrid including Electric Vehicles. *Electronics* **2022**, *11*, 17. https://doi.org/10.3390/ electronics11010017

Academic Editors: Luis Hernández-Callejo, Sergio Nesmachnow and Sara Gallardo Saavedra

Received: 5 December 2021 Accepted: 20 December 2021 Published: 22 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). grid VF magnitude constant, so as to enhance the resilience and increase the penetration of renewable energy to the stand-alone microgrid. In [6], an optimized solution is proposed for minimizing both frequency and voltage deviations. The simultaneous control of VF is achieved with proper load sharing among the DG units. However, the system parameter setting of the traditional control strategy mentioned above is complicated, and the control performance needs to be further improved when faced with complex working conditions such as wind disturbance and load fluctuation.

Therefore, various intelligent algorithms are gradually being widely used in the control of microgrids. In [7], a new scheme for the online minimization of harmonic distortion of an islanded microgrid based on a population-based optimization method is proposed, presenting a new central controller to optimize network voltage harmonics according to particle swarm optimization (PSO) algorithm, while active power is shared between distributed generation units. In [8], a coordinated load shedding control scheme based on Double-*Q* learning for an islanded microgrid is proposed to solve the problem of how to determine the appropriate load shedding amount and objects when frequency is disturbed by considering the relationship between the active power and frequency deviation of each distributed energy resource. However, the intelligent controllers mentioned above can only regulate frequency or voltage, that is, they cannot take both of frequency recovery and voltage adjustment.

Meanwhile, in the construction of the microgrid model, the access of EVs is not considered, and the boundary of the output power of each unit is ignored. Thus, there is room for further optimization in the microgrid model and control strategy. EVs have become a new type of distributed energy storage unit with its energy saving, environmental protection, and flexibility [9,10], which can provide power support for the islanded microgrid and improve its operational flexibility through V2G technology. In [11], an islanded microgrid LFC model including loads, distributed power sources, MT, EVs, and their constraints is established. However, the output power boundary of the EVs charging station model in this paper is a fixed value, which does not match the actual situation. In [12], a microgrid including micro gas turbine (MT), EVs, distributed power, and loads is established, and an improved robust model predictive frequency control strategy of microgrids with EVs is proposed, which can better suppress the frequency fluctuation with a faster response speed than other methods, but the random output power boundary is not refined from the perspective of cluster EVs. In addition, none of the above references considers the reactive power regulation effect of EVs on voltage stability. In fact, the power boundary of the charging station can be affected by user travel demand, charging behavior, and the characteristics of EV clusters. Thus, the active power *P* and reactive power *Q* output by the EVs charging station can be adjusted according to the control command and the power factor angle of the charger, so as to complete the stability control of VF.

In summary, the randomness of users can affect the charging behavior of EVs stations. In addition, there is no suitable intelligent control algorithm that can use EVs to realize the coordinated control of the VF of the islanded microgrid. Thus, a VF coordinated control strategy based on Deep Deterministic Policy Gradient (DDPG) is proposed in this paper, which is applied to the VF control of an islanded microgrid with EVs. The main contributions are as follows: (1) In order to solve the problem of randomness in the charging boundary of EVs caused by the users' randomness, the VF control model of EVs is established. The SOC constraint condition of the EVs is established, and a fourquadrant two-way charger capacity model is introduced. Thus, a microgrid VF control model including load, MT, EVs, and their random power boundary is built; and (2) the voltage and frequency fluctuations can be caused by wind disturbance and load fluctuations. Thus, the DDPG controller with online learning and experience playback capabilities is selected. The convergence characteristics of DDPG are great, so it can coordinate the frequency recovery and voltage regulation of the islanded microgrid greatly. (3) In order to achieve effective regulation of voltage and frequency at the same time, the structure of DDPG controller is designed according to the two control goals of microgrid frequency

and voltage. In addition, then the definition of space, the design of global and local reward functions, and the selection of optimal hyperparameters are completed. Thereby, it can simultaneously meet the VF control requirements.

2. Microgrid Control Model with EVs

The VF control in microgrid can be realized by distributed power supply, energy storage device, etc. In addition, EVs can also participate in microgrid VF regulation. For the VF control of microgrid, the microgrid control system of distributed power supply, load, MT, and EVs is established in this section.

2.1. Electric Vehicle Control Model

As a flexible energy storage device in microgrid control, EVs can regulate the charge and discharge power of the battery according to the instructions of the controller, thereby to control the interaction of active power with the grid [13]. At the same time, charger scheduling is applied to realize the regulation of voltage or reactive power. The two-way charger can realize four-quadrant operation [14], and the power factor cannot determine the transmission direction of reactive power, so the operating quadrant of the charger cannot be determined. Taking the power factor angle as the control variable can determine the transmission direction and magnitude of active and reactive power together, which is more conducive to the two-way transmission control of active and reactive power between the grid and the EV.

The function of EVs in microgrid control is similar to that of energy storage devices. In terms of active power, the charging and discharging power ranges of EV are limited within $\pm \lambda_e$, due to the limits of inverter capacity. The E_{max} is the maximum capacity of EVs station. In addition, the recommended maximum capacity $E_{rmax} = 0.9E_{max}$ and the recommended minimum capacity $E_{rmin} = 0.1E_{max}$ are set to ensure the safe and stable operation of EV station. When the current capacity *E* of the EVs station is higher than the E_{rmax} , the EV stations can discharge to the microgrid, and the discharge power range is $0-\lambda_e$. Similarly, if the current capacity of the EVs station is lower than the E_{rmin} , the EV station can be charged from the microgrid within the charging power range is $-\lambda_e-0$. In addition, the EV control model can be affected by users' uncertain factors such as the randomness of travelling demands and charging behavior of users.

Firstly, the randomness of user travel demand affects the capacity and limitation of the charging station to be random. Therefore, it is necessary to establish the constraints of *SOC* to ensure that the user's normal travel is still satisfied under the interaction between EVs and the grid. In addition, the initial *SOC* of the battery in this paper is set as a random number [15] obeying Gaussian distribution, and its probability density function is expressed as Equation (1):

$$f(s) = \frac{1}{\sigma_s \sqrt{2\pi}} e^{\frac{-(s-\mu_s)^2}{2\sigma_s^2}}$$
(1)

where μ_s represents the average value of *SOC*, and σ_s represents the standard deviation.

j

According to the 2017 National Household Travel Survey (NHTS) of the US Department of Transportation [16], it can be obtained that the daily mileage *L* obeys lognormal distribution, and its probability density function is as follows:

$$f(L) = \frac{1}{L\sigma_{L}\sqrt{2\pi}} e^{\frac{-(lnL-\mu_{L})^{2}}{2\sigma_{L}^{2}}}$$
(2)

where μ_L represents the average value of the daily mileage *L*, and σ_L represents the standard deviation.

According to the daily driving mileage, the charging time T_c is calculated:

$$T_c = \frac{LQ_{100}}{100P_c}$$
(3)

$$I_{\text{leave}} = (1 + \sigma_T) I_c \tag{4}$$

where σ_T is a positive random number.

Based on the above parameters, the demanded *SOC* for future travel named SOC_m can be calculated [17]:

$$SOC_m = S_0 + \frac{L}{L_{\max}} \tag{5}$$

where S_0 is the initial SOC for EVs.

Therefore, for EVs in the station, the *SOC* can be maintained within the range of $[SOC_{rmin}, SOC_{rmax}]$. SOC_{rmax} and SOC_{rmin} are the recommended maximum and minimum value of *SOC*, which can ensure the life of the battery. To satisfy the sufficient SOC_m to make sure the follow-up driving when EVs leave, the constraint conditions are added to the *SOC* of EVs, as shown in Figure 1. The blue dotted line represents the charge boundary, which means that the EV can no longer charge when the *SOC* reaches SOC_{rmax} . The red dotted line represents the discharge boundary, which means that the EV can no longer charge boundary, which means that the SOC reaches SOC_{rmax} . The red dotted line represents the discharge boundary, which means that the EV can no longer charge boundary, which means that the EV can no longer discharge when the *SOC* reaches SOC_{rmin} . The solid green line represents the boundary of forced charging, which means that the EV is forced to charge to ensure the SOC_m when leaving the charging station.



Figure 1. Boundary of charging and discharging constraints for EVs.

Furthermore, in terms of active power, the rated charging power of a single EV can be set to $P_{EV,i}^{ch}$ and the rated discharging power to $P_{EV,i}^{dis}$. The relationship between the charging power of a single EV and the charging and discharging state can be obtained as follows: When $SOC_i \ge SOC_{\text{rmax}}$, the single EV can discharge positive power increment $0 < \Delta P_{EV,i} < P_{EV,i}^{dis}$, which can ensure that SOC_i is controlled below SOC_{rmax} . When $SOC_i \le SOC_{\text{rmax}}$, the single EV can only be charged, that is, only the negative power increment can be discharged $-P_{EV,i}^{ch} < \Delta P_{EV,i} < 0$, which can ensure that SOC_i is controlled above SOC_{rmin} . When $SOC_{\text{rmin}} < SOC_i < SOC_{\text{rmax}}$, the single EV can be charged and discharged. Thus, the power increment satisfies $-P_{EV,i}^{ch} < \Delta P_{EV,i} < P_{EV,i}^{dis}$. In summary, the

instruction distribution of the EVs station through the controller is shown in Figure 2. In addition, the charging and discharging constraint boundary of a single EV can be obtained as follows:

$$P_{EV,i}^{+}(t) = \begin{cases} 0 , t \ge T_{leave,i}, 0 \le t < T_{leave,i} \text{ and } SOC_{EV,i}(t) \le SOC_{EV,i}^{-}(t) \\ P_{EV,i}^{dis}(t), 0 \le t < T_{leave,i} \text{ and } SOC_{EV,i}(t) > SOC_{EV,i}^{-}(t) \end{cases}$$
(6)

$$P_{EV,i}^{-}(t) = \begin{cases} 0 , t \ge T_{leave,i}, 0 \le t < T_{leave,i} \text{ and } SOC_{EV,i}(t) \ge SOC_{EV,i}^{+}(t) \\ P_{EV,i}^{ch}(t), 0 \le t < T_{leave,i} \text{ and } SOC_{EV,i}(t) < SOC_{EV,i}^{+}(t) \end{cases}$$
(7)



Figure 2. The distribution of controller commands in the charging station.

The charging and discharging constraint boundary of the cluster EVs' P_{EV} can be obtained from the boundary of a single EV as follows:

$$\begin{cases}
P_{EV}^{-}(t) < \Delta P_{EV}(t) < P_{EV}^{+}(t) \\
P_{EV}^{+}(t) = \sum_{i=1}^{n_{EV}} P_{EV,i}^{+}(t) \\
P_{EV}^{-}(t) = \sum_{i=1}^{n_{EV}} P_{EV,i}^{-}(t)
\end{cases}$$
(8)

where n_{EV} is the number of EV.

In addition, the active power capacity calculation is related to the number and the *SOC* state of EV:

$$E_{ct} = \sum_{i=1}^{N_{ev}} (SOC_i \times E_i) / E_{all}$$
⁽⁹⁾

where E_i represents the active power capacity of a single EV, E_{all} represents the total active power capacity of EVs, and E_{ct} represents the real time active power capacity of the EVs station.

From this, it can be obtained that the output power ΔP_{EV} of the EV charging station during the charging and discharging process should meet the following constraints:

$$\begin{cases} 0 < \Delta P_{EV} < \lambda_e , & E_{ct} > E_{rmax} \\ -\lambda_e < \Delta P_{EV} < \lambda_e , & E_{rmin} < E_{ct} < E_{rmax} \\ -\lambda_e < \Delta P_{EV} < 0 , & E_{ct} < E_{rmin} \end{cases}$$
(10)

when $E_{ct} > E_{rmax}$, the real time active power capacity E_{ct} of the EV station is higher than the recommended maximum capacity E_{rmax} , due to the rapid increase in the number of EVs in the charging station. When $E_{ct} < E_{rmin}$, the number of EVs in the charging station is too small, or the EVs in the charging station are all in a low battery state. When $E_{rmin} < E_{ct} < E_{rmax}$, the EV station can either discharge to the microgrid or charge from the microgrid.

Furthermore, the capacity state *E* of the EVs station is related to the EVs existing in the EVs station in different *SOC* states. Therefore, by combining Equations (8) and (10), it can obtain the constraint of active output power ΔP_{EV} considering the travel demand of users, the number of electric vehicles, and the real-time *SOC* of electric vehicles as:

$$\begin{cases}
0 < \Delta P_{EV} \le P_{EV}^+(t), & E_{ct} > E_{rmax} \\
P_{EV}^-(t) \le \Delta P_{EV} \le P_{EV}^+(t), & E_{rmin} < E_{ct} < E_{rmax} \\
P_{EV}^-(t) \le \Delta P_{EV} < 0, & E_{ct} < E_{rmin}
\end{cases}$$
(11)

After obtaining the boundary of the active discharge power ΔP_{EV} of the EVs, the reactive power boundary can be obtained through the power factor angle of the charger, and the circuit topology of the four-quadrant bidirectional charger mostly uses a double-buck AC–DC half-bridge conversion circuit, a traditional AC–DC half-bridge conversion circuit, and an AC–DC full-bridge conversion circuit. The capacity curve of the charger is shown in Figure 3 [18].



Figure 3. The capacity curve of the charger.

 φ is the power factor angle when the apparent rated power is ΔS_{EV} . φ_{\min} and φ_{\max} are the minimum and maximum power factor angles of the charger. The positive axis of the *P* axis and *Q* axis represents the energy transferred from the grid to the EV charger. When the active power is *OA*, the adjustable range of reactive power is *CC'*, and the length of *OB* is the apparent rated power ΔS . In addition, the relationship of the active and reactive power ΔP_{EV} and ΔQ_{EV} can be charged by Figure 3, as in the Formula (12):

$$\Delta Q_{EV} = \Delta P_{EV} \tan \varphi \tag{12}$$

$$\begin{cases} \varphi_{\min} < \varphi < \varphi_{\max} \\ -\varphi_{\min} < \varphi < -\varphi_{\max} \end{cases}$$
(13)

Thus, the power factor angle needs to meet the operating characteristics of the charger, and when $\Delta P_{EV} > 0$, the grid feeds active power to the EVs, when $\Delta Q_{EV} > 0$, the grid feeds reactive power to the EVs.

In summary, the boundary of the output power increment of the EV charging station is affected by the number of EV in the charging station N_{EV} , SOC state, electric vehicle charging station real time capacity E, and the angle of charging power factor.

2.2. VF Control Model of Microgrids with EVs

The output characteristics of distributed wind power and photovoltaic system are random, and load fluctuations simultaneously affect the output of active and reactive power. Therefore, in the process of microgrid VF control in this paper, the wind power and photovoltaic system are equivalent to disturbance sources [19]. In addition, the load response characteristics of wind power system and photovoltaic power system are similar, so only the microgrid load VF control under the wind power disturbance is considered, and it is applied using recorded historical data [20]. In addition, the MT is added to the microgrid system as a main control unit in this paper to ensure the flexibility and validity of microgrid regulation.

The structure of the microgrid is in Figure 4. The microgrid includes a MT, EVs, distributed wind power, and load.



Figure 4. Structure of the islanded microgrid.

 ΔP_L and ΔQ_L are the load disturbance power, ΔP_W and ΔQ_W are the wind disturbance power, ΔP_{MT} and ΔQ_{MT} are the power variation of *MT*, and ΔP_{EV} and ΔQ_{EV} are the power variation of EVs.

3. The Design of Microgrid VF Controller Based on DDPG

In the islanded microgrid, it is important to maintain the stability of VF, but there are some control problems such as various uncertainties and nonlinearities caused by DGs and EVs, which can inevitably cause the VF fluctuation and make it deviate from the reference value.

In addition, the Deep Reinforcement Learning (DRL) with online learning, experience playback capabilities and other advantages, is suitable for nonlinear systems [21]. Therefore, in this paper, a VF controller based on DDPG for islanded microgrid with EVs is designed. The frequency and voltage deviation is fed back to the DDPG controller, which adjusts the power output of each unit to ensure the stability of the frequency and voltage of the system.

3.1. Theoretical Analysis of DDPG

Q-learning and Deep Q-learning (DQN) are typical value-based reinforcement learning algorithms that use value functions to learn the optimal strategy during the interaction with the environment [22]. However, since the Q-learning cannot process continuous signals, it is necessary to discretize the action space. Therefore, it is difficult to realize the precise control of MT, EVs and chargers, which is not suitable for the design of this paper.

In addition, the learning of the DDPG can be carried out in a continuous action space [23]. The DDPG contains four networks, namely actor current network, actor target network, critic current network, critic target network. At *t*, the actor current network parameter is θ , and the actor target network parameter is θ' , the critic current network parameter is ω , the critic target network parameter is ω' .

In the above four networks, the actor current network can generate action a_t according to the current status s_t . The actor target network can generate the action a_{t+1} at the t + 1 time according to the subsequent state of the environment. The critic current network can calculate the value R_t corresponding to the status s_t and action a_t . The Critic target network can generate the value of $Q_{value'}(s_{t+1}, a_{t+1} | \omega')$ based on subsequent state s_{t+1} and action a_{t+1} , which is used to calculate the target value y, as shown in the Formula (14):

$$y = r_t + \gamma Q_{value}(s_{t+1}, a_{t+1}, \omega')$$
(14)

where γ is a discount factor and $0 < \gamma < 1$, $Q_{value'}(s_{t+1}, a_{t+1} | \omega')$ is the value generated by subsequent state s_{t+1} and action a_{t+1} , which is used to calculate the target value y.

Meanwhile, the critic current network parameter ω is updated by the gratial direction of the neural network using a mean square difference loss functional Formula (15). In addition, the parameter of the actor current network θ is updated through the gradient of the neural network, as shown in Formula (16):

$$L = \frac{1}{m} \sum_{j=1}^{m} (y_j - Q(s_j, A_j, \omega))^2$$
(15)

$$\nabla J(\theta) = \frac{1}{m} \sum_{j=1}^{m} \left[\left. \nabla_a Q(s, a, \omega) \right|_{s=s_j, a=\pi_\theta(s)} \nabla_\theta \pi_\theta(s) \right|_{s=s_j} \right]$$
(16)

where *m* is the number of samples, y_j is the target value of the *j* sample, $Q(s_j,a_j,\omega)$ is the output value of the critic current network for the *j* sample, and $\pi_{\theta}(\cdot)$ is the output value of the actor current network.

Furthermore, it is necessary to update the critic target network and actor target network parameters by Equation (17):

$$\begin{aligned}
\omega' &\leftarrow \tau \omega + (1 - \tau)\omega' \\
\theta' &\leftarrow \tau \theta + (1 - \tau)\theta'
\end{aligned}$$
(17)

where τ is an update coefficient, which is generally small.

In addition, the *E* is a termination function, which is to determine whether the Agent enters the termination. If the Agent enters the termination state, the iterative process stops and a new round of state sequence starts. If the Agent enters the non-termination state, the iterative process of the wheel can be continued.

In summary, status information, reward value, action information, and termination status information {*s*, *a*, *R*, *s'*, *E*} are formed into a sample unit and stored in the empirical playback set *D*. Then, *m* sample units of set *D* are taken to be trained by Formulas (14)–(17). A total of *T* rounds is trained, and the training step length of each round is T_m . The specific training process is shown in Figure 5.



Figure 5. Structure of the islandeded microgrid.

3.2. Design of DDPG VF Controller Structure

Considering MT and EV output power increment limiting constraints, a VF controller structure based on DDPG is proposed, as shown in Figure 6. The controller is composed of two layers: coordinate layer and control layer. The coordinate layer provides real-time regulation signal ΔA to the control layer according to the frequency deviation Δf , voltage deviation ΔU , and the real-time boundary of output power of EV charging station, and then controls the output power of MT and EV to quickly suppress the frequency and voltage deviation.



Figure 6. Microgrid LFC controller structure based on DDPG.

3.3. Definition of Space and Reward Function

As mentioned above, the state set of the control system is frequency deviation $\Delta F(t)$, voltage deviation $\Delta U(t)$, and the real-time boundary of output power of EV charging station $P_{EV}^{\pm}(t)$ and $Q_{EV}^{\pm}(t)$, so the state space *S* can be defined as follows:

$$S = [\Delta F(t), \Delta U(t), P_{EV}^{+}(t), P_{EV}^{-}(t), Q_{EV}^{+}(t), Q_{EV}^{-}(t)]$$
(18)

In addition, the joint action set *A* of the DDPG controller, namely the output of the controller, should be a real-time set of dispatch instruction of the active and reactive power output of MT, the output active power of EVs, and the power factor angle of the charger. Thus, the action space *A* can be defined as follows:

$$A = [\Delta A_{P,\text{MT}}(t), \Delta A_{P,EV}(t), \Delta A_{Q,MT}(t), \Delta A_{\varphi,EV}(t)]$$
(19)

In addition, then, China's power safety work principle stipulates that the frequency of the power system during normal operation should be within the range of 50 ± 0.2 Hz, and the voltage deviation should within 5%. Thus, on this basis, a certain adjustment dead zone is considered, the discrete set of real-time frequency deviation $\Delta F(t)$ can be set as $(-\infty, -0.2)$, [-0.2, -0.15), [-0.15, -0.10), [-0.10, -0.03), [-0.03, 0.03], (0.03, 0.10], (0.10, 0.15], (0.15, 0.2], $(0.2, +\infty)$, unit of Hz, and the discrete set of real-time voltage deviation $\Delta U(t)$ can be set as (-1, -0.05), [-0.05, -0.03), [-0.03, -0.02), [-0.02, -0.01), [-0.01, 0.01], (0.01, 0.02], (0.02, 0.03], (0.03, 0.05], (0.05, 1), unit of p.u.

Meanwhile, the control objectives in this paper are: ①Restore the frequency to the rated value; ②Regulate and control the voltage to restore to the best state. As a result,

a comprehensive reward function including two local reward functions can be set up to coordinate frequency recovery and voltage adjustment:

$$R = r_f + r_u \tag{20}$$

$$r_{f} = \begin{cases} 0 & |\Delta f| < 0.03 \\ -\mu_{1} |\Delta f| & 0.03 \le |\Delta f| < 0.10 \\ -\mu_{2} |\Delta f| & 0.10 \le |\Delta f| < 0.15 \\ -\mu_{3} |\Delta f| & 0.15 \le |\Delta f| < 0.2 \\ -\mu_{4} |\Delta f| & 0.2 \le |\Delta f| \end{cases}$$
(21)
$$r_{\mu} = \begin{cases} 0 & |\Delta U| < 0.01 \\ -\delta_{1} |\Delta U| & 0.01 \le |\Delta U| < 0.02 \\ -\delta_{2} |\Delta U| & 0.02 \le |\Delta U| < 0.03 \end{cases}$$
(22)

$$\left\{ \begin{array}{c} -\delta_2 |\Delta U| & 0.02 \leq |\Delta U| < 0.05 \\ -\delta_3 |\Delta U| & 0.03 \leq |\Delta U| < 0.05 \\ -\delta_4 |\Delta U| & 0.05 \leq |\Delta U| < 1 \end{array} \right.$$

where *R* is the global reward, r_f is the frequency reward, r_u is the voltage reward, μ_1 , μ_2 , μ_3 and μ_4 are the weights corresponding to the reward function of each control region in the frequency penalty item r_f , and δ_1 , δ_2 , δ_3 and δ_4 and are the weights corresponding to the voltage control regions.

The control process needs to control the frequency through r_f , when $|\Delta f|$ is in adjusting dead zone [-0.05, 0.05] Hz, and the frequency meets the minimum error requirement of normal operation, so the maximum reward value given to the DDPG controller at this time is 0. When $|\Delta f|$ is respectively in normal control (0.05, 0.10) and (0.10, 0.15) Hz, auxiliary control area (0.15, 0.2) Hz, emergency control area (0.2, + ∞) Hz, the controller can get the corresponding negative incentives, namely the penalty value. Meanwhile, when voltage control is performed, the voltage needs to be regulated by r_u , when $|\Delta U|$ is in adjusting dead zone [-0.01, 0.01], the maximum reward value given to the DDPG controller at this time is 0, and when $|\Delta U|$ is respectively in normal control (0.01, 0.02) and (0.02, 0.03), auxiliary control area (0.03, 0.05), emergency control area (0.05, 1), the controller can get the corresponding penalty value.

When determining the values of the above parameters, it should be noted that the size of the reward value can affect the convergence effect and the learning speed. Therefore, it is necessary to perform simulation tests based on actual calculation examples, and the specific process will be discussed later.

In summary, the state space and reward function designed in this paper can realize the simultaneous adjustment of voltage and frequency. When the frequency is restored, it can consider whether the voltage exceeds the limit, and, when adjusting the voltage, it can also consider whether the frequency deviates from the rated value, which significantly improves the overall stability of the microgrid.

3.4. The Selection of Hyperparameter

In DRL, it is necessary to provide the agent with a set of optimal hyperparameters to improve the performance and effect of learning [24].

First of all, the larger the discount factor γ , the more the agent attaches importance to past experience and can give up current interests and pursue overall interests. However, if γ is too large, it will also cause the training of agent to fail to converge. The greater the learning rate α , the faster the agent converges, but the worse the stability; the smaller the α , the better the stability, but the slower the agent converges. Therefore, the convergence speed should be improved on the premise when the agent training can converge. In addition, the design of network structure can be discussed from two aspects: network type and network depth. The choice of network type depends largely on the state space, and the state space of the control system in this paper is frequency and voltage deviation, which belong to

one-dimensional vector, so the full connection layer can better meet the requirements of the storage strategy set. In addition, the network depth determines the generalization ability of the neural network, which includes the number of layers of the neural network h and the neurons in each layer u.

In addition, the specific values of γ , α , h and u need to be selected according to the calculation example.

3.5. Summary of Control Strategy

In summary, the control strategy of this paper is carried out in the following steps:

- 1. Firstly, definite the state set of the control system as $\Delta F(t)$, $\Delta U(t)$, $P_{EV}^{\pm}(t)$ and $Q_{EV}^{\pm}(t)$. In addition, the action space can be defined as $\Delta A_{P,MT}(t)$, $\Delta A_{P,EV}(t)$, $\Delta A_{Q,MT}(t)$, $\Delta A_{\varphi,EV}(t)$.
- 2. Secondly, the parameters are adjusted according to the actual calculation example, and the values of the reward function coefficients and hyperparameters are obtained.
- 3. Thirdly, perform agent training according to the process in Figure 5, and obtain the optimal value function Q network $Q_{\varphi(s,a)}$.
- 4. Finally, in different cases, input disturbances to the islanded microgrid system, and the agent can generate corresponding actions based on the disturbances to adjust the output of each unit, so as to ensure the frequency and voltage balance of the islanded microgrid system.

4. Simulation Results

In order to evaluate the control effect of the above strategy, the coupled islanded microgrid system is built as shown in Figure 7. In addition, the specific settings of equipment parameters are shown in Table 1. The verification of the calculation examples in this paper is carried out through simulation experiments. The computing platform is a PC with i7-1165G7@2.80GHz CPU and 16 GB RAM, and the software environment is Windows 10 Professional and MATLAB R2021a.



Figure 7. Microgrid LFC controller structure based on DDPG.

In the microgrid, there is a MT with capacity of 40 kW, a WT with capacity of 20 kW, an EV station1 with capacity of 16 kW, an EV station2 with capacity of 14 kW, and 60 kW ordinary loads. In addition, this paper assumes that the initial state of the microgrid is stable. Thus, when there is no external disturbance, the power output of MT, EV stations,

WT, and conventional loads are always in balance. Therefore, in the following calculation examples, only the per-unit value of the power fluctuations of MT, EVs stations, WT, and load need to be considered.

Unit	Parameter	Meaning	Value
	T_{f}	time constant of governor	10 s
	T_t	time constant of generator	0.1 s
	R_f	speed regulation factor	0.005 Hz/p.u.
MT	λ_{Pmtd}	lower limit of active power variation	−0.025 p.u.
	λ_{Pmtp}	upper limit of active power variation	0.025 p.u.
	λ_{Qmtd}	lower limit of reactive power variation	−0.025 p.u.
	λ_{Qmtp}	upper limit of reactive power variation	0.03 p.u.
EV1	T_{e1}	time constant of EV	1 s
	λ_{Ped1}	lower limit of active power variation	−0.016 p.u.
	λ_{Pep1}	upper limit of active power variation	0.016 p.u.
	λ_{Qed1}	lower limit of reactive power variation	−0.015 p.u.
	λ_{Qep1}	upper limit of reactive power variation	0.015 p.u.
	n_{EV1}	Initial number of EVs in station1	40
EV2	T_{e2}	time constant of EV	1.5 s
	λ_{Ped2}	lower limit of active power variation	-0.014 p.u.
	λ_{Pep2}	upper limit of active power variation	0.014 p.u.
	λ_{Qed2}	lower limit of reactive power variation	−0.0135 p.u.
	λ_{Qep2}	upper limit of reactive power variation	0.0135 p.u.
	n_{EV2}	Initial number of EVs in station2	35

Table 1. Parameters of equipment in microgrids.

4.1. Pre-Learning Stage

Before the controller is used, it needs to undergo a random trial and error learning process, which is called the pre-learning stage. In the initial stage of pre-learning, the controller has not accumulated any experience and has no intelligent control ability [25]. Only after accepting various state actions can the optimal value function Q network $Q_{\varphi(s,a)}$. Therefore, the wind and load disturbances superimposed by various different amplitudes and different types of functions are set up for repeated training of the controller. Meanwhile, according to the output capacity change data of the electric vehicle charging station, a boundary function of the output power increment that changes randomly over time is set. Take active power disturbance and the output boundary of the active power of EVs as examples. The random disturbance of a certain training process is shown in Figure 8.



Figure 8. Random perturbation function in the pre-learning phase: (**a**) Random Function of Active Power Disturbance; (**b**) Random Function of EV Output Power Boundary.

Meanwhile, through a large number of simulation studies, μ_1 , μ_2 , μ_3 , and μ_4 are referred as 1, 5, 10, and 20, respectively, δ_1 , δ_2 , δ_3 and δ_4 are referred as 5, 20, 50, and

100 respectively, and α and γ are referred as 0.01, 0.09. Meanwhile, the number of learning iterations of the DDPG controller is set to 500, each with 500 steps, and the step length is 0.1 s. Therefore, six groups of parameters (*h*, *u*) are set for the convergence test, and the learning results are shown in Table 2. It can be seen that the reward value of the system at convergence is the highest when *h* = 5 and *u* = 50.

SN	Parameter Settings	Average Reward	Final Award
1	h = 3, u = 50	-34.037	-1.83622
2	h = 3, u = 200	-26.673	-1.20923
3	h = 5, u = 50	-21.096	-0.65307
4	h = 5, u = 200	-27.075	-1.35723
5	h = 10, u = 50	-34.572	-2.54778
6	h = 10, u = 200	-40.922	-3.04643

Table 2. Convergence test results under different parameters.





Figure 9. The complete trend graph of the reward function: (**a**) The complete trend graph of the reward function; (**b**) the trend graph of the reward function for the last 50 iterations.

It can be seen that the agent basically converges after 80 iterations, and the system judges that the learning process has been completed and stops the training after 248 iterations. In this case, the average reward is -21.096 and the final award is 0.65307, which shows that the controller can complete the subsequent simulation at this time.

4.2. The Implementation of Constraint Conditions in the EV Model

In order to verify the implementation of constraint conditions in the EV model, this paper selects several typical monomer EV *SOC* simulation situations as examples, as shown in Figures 10 and 11. In addition, to ensure the life of battery, the initial *SOC* is set between $SOC_{rmin} = 0.2$ and $SOC_{rmax} = 0.8$.

The first situation in Figure 10 shows that, when $SOC < SOC_{rmin}$, the EV will be forced to enter the charging state. Only when $SOC > SOC_{rmin}$ can the EV participate in system regulation. The second situation in Figure 10 shows that, when the EV is close to the leaving time and $SOC < SOC_m$, it will turn to the forced charging state to ensure that the SOC reaches the expected SOC_m when leaving the charging station. In general, the changes in the SOC of EVs participating in the regulation of the microgrid are shown in Figure 11. The SOC of EVs will change in the constraint range.

4.3. Case Study

After completing the pre-learning phase and the verification of the EV SOC constraints, the example can be simulated under different operation scenarios. Meanwhile, in order to



evaluate the effect of DDPG controller proposed in this paper, traditional PID controller and $R(\lambda)$ controller are used in the same scene respectively, and the corresponding controller parameters are shown in Table 3.

Figure 10. Changes of SOC of EVs at critical value.



Figure 11. Changes of SOC of EVs in the normal range.

Controller	Describe	Parameter	Value
PID	Proportional gain	K _P	4
	Integral gain	K_i	1.18
	Differential gain	K _D	0.5
R(λ)	learning rate	α	0.01
	discount factor	γ	0.9
	network depth	(<i>h</i> , <i>u</i>)	(3, 10)

Table 3. PID and $R(\lambda)$ controller system parameters.

4.3.1. Case 1: The Response of Wind Power Disturbance

First of all, wind power disturbance is added to the islanded microgrid system, and wind mainly provides active power disturbances to the grid. In order to compare the adjusting speed of each controller, the wind power disturbance ends after 43 s. The disturbance setting is shown in Figure 12.



Figure 12. Wind power disturbance.

There is not the fluctuation of reactive power in this case, so the impact of voltage fluctuation is not considered here. The variation of frequency deviation under wind power disturbance is shown in Figure 13. Meanwhile, according to the simulation results, this paper takes the absolute value of $|\Delta f|$ as the evaluation object, and sets the threshold of the frequency deviation excellence rate to 2×10^{-4} Hz, and defines T_{recover} as the time which is taken for $|\Delta f|$ to recover to 5×10^{-5} Hz after the wind power disturbance ends. The results of the control test under wind disturbance are shown in Table 4.



Figure 13. Performance of frequency control under wind power disturbance.

Indicators	PID	R (λ)	DDPG
Average (Hz)	$0.9215 imes 10^{-3}$	$0.1159 imes 10^{-3}$	$0.0267 imes10^{-3}$
Maximum (Hz)	$7.174 imes10^{-3}$	$1.148 imes 10^{-3}$	$1.133 imes 10^{-3}$
Proficiency (%)	30.56%	92.07%	98.35%
$T_{\rm recover}$ (s)	34.25	23.95	0.75

Table 4. Frequency simulation results under wind disturbance.

It can be seen from Figure 13 and Table 4 that, compared with the PID controller, the DDPG and $R(\lambda)$ controller with the ability of online learning and experience playback can more effectively deal with the highly random disturbance. Under the wind disturbance, the frequency fluctuation of the islanded microgrid under the DDPG controller can be limited in 2 × 10⁻⁴ Hz, and the excellent rate can reach 98%, which is significantly better than the traditional controller. In addition, if only analyzed from the perspective of frequency control, the control strategy of DDPG and $R(\lambda)$ controller in this paper possesses virtues of great control effect, smaller amplitude of frequency fluctuation, and faster regulation speed than a traditional controller. Furthermore, the regulation speed of DDPG controller is much faster than a $R(\lambda)$ controller.

Furthermore, the power variations of each equipment in islanded microgrid under the DDPG controller are shown in Figure 14. It can be seen that, when the system suffers disturbance, the MT undertakes the main work of frequency regulation, and the output power of EV charging station is also significant. In addition, when the limit is reached, the power variations of different charging stations are different.



Figure 14. Power variations of each equipment under wind power disturbance.

4.3.2. Case 2: The Response to Load Power Disturbance

The fluctuation of load power is gentler than that of wind power, but the load change is abrupt and can cause fluctuations in active and reactive power at the same time. In this case, load power variations are set as $\Delta P_{\rm L} = -0.025$ p.u during 10–40 s, $\Delta P_{\rm L} = 0.005$ p.u during 40–55 s, $\Delta P_{\rm L} = -0.0025$ p.u during 55–70 s, $\Delta P_{\rm L} = 0.015$ p.u during 70–150 s, $\Delta Q_{\rm L} = -0.04$ p.u during 7.5–34 s, $\Delta Q_{\rm L} = -0.015$ p.u during 34–66 s, $\Delta Q_{\rm L} = 0.005$ p.u during 66–96 s, $\Delta Q_{\rm L} = -0.0075$ p.u during 96–150 s. The specific setting of load disturbance is shown in Figure 15.



Figure 15. Load power disturbance: (**a**) load active power disturbance; (**b**) load reactive power disturbance.

The DDPG controller is compared with traditional PID and $R(\lambda)$ controller, and the frequency and voltage fluctuation are shown in Figures 16 and 17. The same as the case 1, this part takes $|\Delta f|$ and $|\Delta U|$ as the evaluation object, and sets the threshold of the $|\Delta f|$ excellence rate to 2×10^{-4} Hz, the $|\Delta U|$ excellence rate to 0.01 p.u. Meanwhile, T_{recover} is defined as the time which is taken for $|\Delta f|$ to recover to 5×10^{-5} Hz and $|\Delta U|$ to recover to 0.002 p.u after the load power disturbance no longer changes. Thus, the statistical results of the control test under load disturbance are shown in Tables 5 and 6.



Figure 16. Performance of frequency control under load power disturbance.



Figure 17. Performance of voltage control under load power disturbance.

Indicators	PID	R (λ)	DDPG
Average (Hz)	$0.995 imes 10^{-3}$	$0.437 imes10^{-3}$	$0.0521 imes 10^{-3}$
Maximum (Hz)	$3.54 imes10^{-3}$	$1.63 imes 10^{-3}$	$1.09 imes 10^{-3}$
Proficiency (%)	9.6%	16.7%	98%
$T_{\rm recover}$ (s)	/	/	0.13

Table 5. Frequency simulation results under load disturbance.

Table 6. Voltage simulation results under load disturbance.

Indicators	סוק	$\mathbf{P}(\lambda)$	DDPC
Indicators	ПD	K (<i>N</i>)	DDIG
Average (p.u)	0.0093	0.0021	0.00047
Maximum (p.u)	0.0569	0.0112	0.0023
Proficiency (%)	30.3%	83.07%	100%
$T_{\rm recover}$ (s)	/	/	9.2

It can be seen from Figures 16 and 17 and Tables 5 and 6 that, when the load changes, compared with the PI controller and $R(\lambda)$ controller, the DDPG controller can ensure that the frequency deviation of the microgrid is maintained within $\pm 1 \times 10^{-3}$ Hz Hz, and the voltage deviation is also close to 0, which is much smaller than the control index of the power quality of the power grid. In addition, compared with the $R(\lambda)$ controller, the DDPG controller, the DDPG controller can coordinate the frequency recovery and voltage adjustment of the islanded microgrid, so as to meet the VF control requirements at the same time, which has superior dynamic control characteristics.

Furthermore, the power variations of each equipment are shown in Figure 18. The MT in the micro grid is used as the main source to maintain the stability of the VF amplitude of the microgrid, while the EV_1 and EV_2 as the slave sources are mainly responsible for the regulation of the active power of the microgrid and also participate in the regulation of the reactive power. In addition, due to the randomness of users, the output power boundary of EV charging stations is random, showing obvious jagged shapes.





5. Conclusions

To solve the problem in which the stability of island microgrid is greatly affected by random power sources, and it is difficult to control frequency and voltage together, a VF control strategy of islanded microgrids with EVs is proposed in this paper. The randomness of charging behavior is considered, and an islanded microgrid system including MT, WT,

EVs stations, and loads is established. Thus, a VF synergistic control strategy based on DDPG is proposed. The simulation results show that:

- 1. Compared with PID controller, the DDPG controller with the ability of online learning and experience playback can more effectively deal with the highly random disturbance. Under the wind disturbance, the frequency fluctuation of the islanded microgrid under the DDPG controller can be limited in 2×10^{-4} Hz, and the excellent rate can reach 98%, which is significantly better than the traditional controller.
- 2. Compared with the $R(\lambda)$ controller, the DDPG controller in this paper can coordinate the frequency recovery and voltage adjustment of the island microgrid, so as to meet the VF control requirements at the same time, which is more suitable for the stable control of the microgrid. When the load changes, the DDPG controller can ensure that the frequency deviation of the microgrid is maintained within $\pm 1 \times 10^{-3}$ Hz, and the voltage deviation is also close to 0.
- 3. The EV charging station has the characteristics of small inertia and fast regulation speed in the microgrid control, which can play an important role in VF regulation;
- 4. The realization effect of the constraint conditions in the EV model is great. The single EV can judge whether it participates in the adjustment of the microgrid system according to the *SOC* situation.

For microgrid systems with more complex structures and larger volumes, it is necessary to consider the multi-microgrid interconnection technology. In addition, multi-agent algorithms such as MA-DDPG, COMA, CommNet, etc. will also be applied to the control of multi-microgrid. The follow-up work will focus on in-depth analysis and research in these directions, and add corresponding hardware circuit experiments or semi-physical simulation experiments.

Author Contributions: P.F., S.K. (Song Ke) and S.K. (Salah Kamel) conceptualized the idea of this research, P.F.; performed the experiments and data analysis, P.F. and S.K. (Song Ke) wrote the paper; J.Y., Y.L., J.X., B.X. and G.I.R. provided supervision and reviewed the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the science and technology project of the State Grid Corporation of China Research and the application of flexible control technology for a distribution system with large-scale distributed generation and a multi microgrid (No. 52093220000H).

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

- 1. Lam, Q.L.; Bratcu, A.I.; Riu, D.; Boudinet, C.; Labonne, A.; Thomas, M. Primary frequency H_∞ control in stand-alone mi-crogrids with storage units: A robustness analysis confirmed by real-time experiments. *Int. J. Electr. Power Energy Syst.* 2020, 115, 105507.1–105507.13. [CrossRef]
- Chae, S.; Kim, G.; Choi, Y.-J.; Kim, E.-H. Design of Isolated Microgrid System Considering Controllable EV Charging Demand. Sustainability 2020, 12, 9746. [CrossRef]
- Ziras, C.; Prostejovsky, A.M.; Bindner, H.W.; Marinelli, M. Decentralized and discretized control for storage systems offering primary frequency control. *Electr. Power Syst. Res.* 2019, 177, 106000.1–106000.10. [CrossRef]
- 4. Roudbari, E.S.; Beheshti, M.T.H.; Rakhtala, M. Voltage and frequency regulation in islandeded microgrid with PEM fuel cell based on a fuzzy logic voltage control and adaptive droop control. *IET Power Electron.* **2019**, *13*, 78–85. [CrossRef]
- Joung, K.W.; Kim, T.; Park, J.W. Decoupled frequency and voltage control for stand-alone microgrid with high renewable penetration. In Proceedings of the IEEE/IAS Industrial & Commercial Power Systems Technical Conference IEEE, Niagara Falls, ON, Canada, 7–10 May 2018.
- Dhua, R.; Goswami, S.K.; Chatterjee, D. An Optimized Frequency and Voltage Control Scheme for Distributed Generation Units of an Islandeded Microgrid. In Proceedings of the 2021 Innovations in Energy Management and Renewable Resources (IEMRE), Kolkata, India, 5–7 February 2021.
- 7. Keypour, R.; Adineh, B.; Khooban, M.H.; Blaabjerg, F. A New Population-Based Optimization Method for Online Minimiza-tion of Voltage Harmonics in Islandeded Microgrids. *IEEE Trans. Circuits Syst. II: Express Briefs* **2019**, *67*, 1084–1088. [CrossRef]

- Wang, C.; Mei, S.; Dong, Q.; Chen, R.; Zhu, B. Coordinated Load Shedding Control Scheme for Recovering Frequency in Is-landeded Microgrids. *IEEE Access* 2020, *8*, 215388–215398. [CrossRef]
- 9. Iqbal, S.; Xin, A.; Jan, M.U.; Salman, S.; Abdelbaky, M.A. V2G Strategy for Primary Frequency Control of an Industrial Mi-crogrid Considering the Charging Station Operator. *Electronics* **2020**, *9*, 549. [CrossRef]
- 10. Fan, H.; Lin, J.; Zhang, C.K.; Mao, C. Frequency regulation of multi-area power systems with plug-in electric vehicles con-sidering communication delays. *IET Gener. Transm. Distrib.* 2016, *10*, 3481–3491. [CrossRef]
- 11. Yang, J.; Zeng, Z.; Tang, Y.; He, H.; Wu, Y. Load Frequency Control in Isolated Micro-Grids with Electrical Vehicles Based on Multivariable Generalized Predictive Theory. *Energies* **2015**, *8*, 2145–2164. [CrossRef]
- 12. Rao, Y.; Yang, J.; Xiao, J.; Xu, B.; Liu, W.; Li, Y. A frequency control strategy for multi-microgrids with V2G based on the im-proved robust model predictive control. *Energy* **2021**, 222, 119963.1–119963.13. [CrossRef]
- 13. Li, P.; Hu, W.; Xu, X.; Huang, Q.; Liu, Z.; Chen, Z. A frequency control strategy of electric vehicles in microgrid using virtual synchronous generator control. *Energy* **2019**, *189*, 116389. [CrossRef]
- Kisacikoglu, M.C.; Ozpineci, B.; Tolbert, L. EV/PHEV Bidirectional Charger Assessment for V2G Reactive Power Operation. *IEEE Trans. Power Electron.* 2013, 28, 5717–5727. [CrossRef]
- 15. Cao, Y.; Tang, S.; Li, C.; Zhang, P.; Tan, Y.; Zhang, Z.; Li, J. An Optimized EV Charging Model Considering TOU Price and SOC Curve. *IEEE Trans. Smart Grid* 2011, *3*, 388–393. [CrossRef]
- U.S. U.S. Department of Transportation. Federal Highway Administration, 2017 National Household Travel Survey. Available online: http://nhts.ornl.gov (accessed on 20 November 2021).
- 17. Gjelaj, M.; Hashemi, S.; Andersen, P.B.; Traeholt, C. Optimal infrastructure planning for EV fast-charging stations based on prediction of user behaviour. *IET Electr. Syst. Transp.* **2020**, *10*, 1–12. [CrossRef]
- Mojdehi, M.N.; Ghosh, P. An On-Demand Compensation Function for an EV as a Reactive Power Service Provider. *IEEE Trans. Veh. Technol.* 2015, 65, 4572–4583. [CrossRef]
- 19. Vachirasricirikul, S.; Ngamroo, I. Robust LFC in a Smart Grid with Wind Power Penetration by Coordinated V2G Control and Frequency Controller. *IEEE Trans. Smart Grid* 2014, *5*, 371–380. [CrossRef]
- Zhu, X.; Xia, M.; Chiang, H.-D. Coordinated sectional droop charging control for EV aggregator enhancing frequency stability of microgrid with high penetration of renewable energy sources. *Appl. Energy* 2018, 210, 936–943. [CrossRef]
- Zhang, J.; Zhang, C.; Chien, W.C. Overview of Deep Reinforcement Learning Improvements and Applications. J. Internet Technol. 2021, 22, 239–255.
- Giannopoulos, A.; Spantideas, S.; Kapsalis, N.; Karkazis, P.; Trakadas, P. Deep Reinforcement Learning for Energy-Efficient Multi-Channel Transmissions in 5G Cognitive HetNets: Centralized, Decentralized and Transfer Learning Based Solutions. *IEEE* Access 2021, 9, 129358–129374. [CrossRef]
- Yang, Q.; Zhu, Y.; Zhang, J.; Qiao, S.; Liu, J. UAV Air Combat Autonomous Maneuver Decision Based on DDPG Algorithm. In Proceedings of the 2019 IEEE 15th International Conference on Control and Automation (ICCA) IEEE, Edinburgh, Scotland, 16–19 July 2019.
- 24. Gulde, R.; Tuscher, M.; Csiszar, A.; Riedel, O.; Verl, A. Deep reinforcement learning using cyclical learning rates. In Proceedings of the 2020 Third International Conference on Artificial Intelligence for Industries (AI4I), Irvine, CA, USA, 21–23 September 2020.
- 25. Yu, T.; Zhou, B.; Chan, K.W.; Yuan, Y.; Yang, B.; Wu, Q.H. R(λ) imitation learning for automatic generation control of interconnected power grids. *Automatica* **2012**, *48*, 2130–2136. [CrossRef]