IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 3 June 2024, accepted 24 July 2024, date of publication 2 August 2024, date of current version 13 August 2024.

*Digital Object Identifier 10.1109/ACCESS.2024.3437746*

RESEARCH ARTICLE

# Auxiliary Decision Method for Power Dispatching Based on Flexible Super-Capacitors and Proximal Policy Optimization Algorithm

**SHENG YANG, JINGLONG HE, AND JINMING LIU**
Power Dispatching Control Center, Guangxi Power Grid, Nanning 535000, China
Corresponding author: Sheng Yang (ys202403@163.com)

**ABSTRACT** To reduce power waste and improve the operational safety of renewable energy grids, this study proposes an auxiliary decision-making method for power dispatching based on flexible super-capacitors and proximal policy optimization algorithms. The results demonstrated that the impedance of Mn&Ni-based composite electrodes was small, about 118 Ω, and when the current density was 1A/g, the constant current charging and discharging time could reach 128 seconds, with a specific capacitance of 39F/g. In addition, when the power density of the capacitor was 8500W/kg, the energy density could reach up to 125Wh/kg. As for the auxiliary decision-making model for power dispatching, the reward value of the proximal policy optimization algorithm began to converge after approximately 50,000 iterations, at which point its reward value was approximately 257. The reward value of the one-step scheduling strategy gradually stabilized at around 1.0 after the number of scheduling steps reached 560,000. When the energy output rate of renewable energy units was high, the consumption rate of renewable energy was 90.1%. The above results indicate that the power scheduling method combining flexible super-capacitors and proximal policy optimization algorithms can effectively achieve the storage of electricity and improve the consumption rate of renewable energy.

**INDEX TERMS** Flexible super-capacitors, composite materials, proximal policy optimization, power dispatch, assisting decision-making.

## I. INTRODUCTION

The current energy structure of the world is still dominated by Fossil Fuels (FFs). According to statistics, the total annual energy consumption worldwide is about 13.4 billion tons of standard coal, of which FFs account for 85%, and most of the electricity also relies on FFs. The above situation has also led to the depletion of FFs and exacerbated pollution. Therefore, to alleviate the energy crisis and protect the ecological environment, it is crucial to increase the utilization of Renewable Energy (RE). The use of the RE power grid can greatly alleviate the dependence of electricity on FFs and help alleviate the problem of energy scarcity [1], [2]. However, RE owns strong volatility and randomness, and the generated electricity relies on power electronic devices connected to the grid. The low anti-interference, weak support, and zero inertia features of power equipment make the safety and stability of the power grid increasingly prominent. The accumulation of quantitative changes can lead to a qualitative change in the characteristics of the power grid operation [3]. Therefore, it is necessary to develop a scheduling complex decision-making method suitable for RE power grid to alleviate its impact on the safety of power grid operation. Tang H et al. put forward a scheduling strategy built on Deep Q-learning (D-QL) for the dynamic optimization scheduling problem of ultra-short-term wind-power prediction. This method first established a dynamic optimization scheduling model with the target of minimizing the operating cost of the basic power system and transformed the optimal scheduling problem into a Markov

The associate editor coordinating the review of this manuscript and approving it for publication was Wencong Su.

decision process. Finally, this method used D-QL to optimize the scheduling strategy, which had high feasibility [4]. Huang Y et al. designed a scheduling method grounded on improved stochastic dual dynamic integer programming for multi-stage distributed robust power generation scheduling optimization problems. This method fully considered the independent uncertainty of wind power and conventional loads, and combined the advantages of multi-stage DRO models and non-competitiveness into power generation scheduling. This method had better scheduling performance compared to other methods [5]. Anandam PV et al. established a multi-objective reactive power management method that considers the uncertainty of real-time generation and load for reactive power management in power systems. This method utilized index-based evolutionary techniques to schedule reactive power and extracted the optimal compromise solution of the scheduling scheme using an ''S-shaped'' fuzzy membership function. Its scheduling performance has been proven to be superior to other methods [6].

Although effective scheduling of RE grids can reduce power differentials, there is still a significant amount of electricity wasted. If the aforementioned electricity can be stored, it can effectively alleviate the problem of energy waste. Flexible Super-capacitors (FSCs) have the advantages of high capacitance, long lifespan, and high Energy Density (ED), making them suitable as energy storage devices in the power grid to reduce power waste. Ojeda L et al. proposed a preparation method on the grounds of expired drugs to address the issue of FSCs indicators. This method used expired drug powders of diclofenac and ibuprofen deposited on the electrodes of super capacitors to assemble capacitors. After adding drug powder, the capacitance of the capacitor increased by 2.5-6.3 times, and the highest ED value could reach 103.5Wh/kg [7]. Macherla N et al. proposed a nanocomposite material based on Naphthalene Sulfonic Acid (NSA) doped Polyaniline (PANI)/sulfur-doped reduced graphene oxide (RGO) to improve the performance of FSCs. This material was grown and arranged on sulfur-doped RGO nanosheets, using NSA as a doping agent for PANI and soft templates. At a Current Density (CD) of 1A/g, the Specific Capacitance (SC) of this material was as high as 347.5F/g, and after 2500 cycles, its initial capacitance only lost 11% [8]. Wang S et al. proposed an electrolyte based on MXene crosslinked organic gel to optimize the electrolyte performance of FSCs. The MXene polyacrylic acid N-hydroxyethyl acrylamide hydrogel network of the electrolyte was replaced by a mixture of LiCI and ethylene glycol solvent. This electrolyte had good conductivity and operated at a temperature of −20°C-80°C [9].

In summary, current research on the scheduling of RE power grids and FSCs has been quite effective, but there are relatively few studies on the simultaneous application of both in the power grid. Research in this field often focuses on one aspect while neglecting the interaction between the two. Therefore, to improve the storage and utilization efficiency of electrical energy in the power grid, achieve scheduling optimization of renewable power grids, and reduce power differences in the power grid, this paper proposes a power grid scheduling Assisted Decision-Making (ADM) method based on FCS and Proximal Policy Optimization (PPO) algorithm. This method innovatively constructs FSCs based on Mn&Ni-based Composite Electrode Materials (Mn&Ni-CEM) and introduces them into the power grid. At the same time, the PPO algorithm is used to optimize the scheduling of the power grid to achieve a balance between the two.

## II. METHODS AND MATERIALS

In response to the problem of energy waste in high proportion RE power systems, this study proposes a power storage method based on FSCs and a power dispatch ADM model based on PPO algorithm to reduce power deviation in the power system and ensure safe and economical operation.

### A. PREPARATION OF Mn&Ni-CEM-BASED FSCs

Experimental materials and equipment. FSCs are composed of flexible substrates, Electrode Materials (EMs), and solid electrolytes. The EM can simultaneously store energy and collect current, while the solid electrolyte can simultaneously act as an electrolyte and a separator. Compared with traditional super-capacitors, FSCs are good at safety, lightweight, low production cost, and environmental protection. To improve the energy storage performance of FSCs, this study proposes FSCs based on Mn&Ni-CEM, and the required materials for their preparation are shown in Table 1.

In Table 1, the required reagents include AcB (particle size: 35-45 nm), polytetrafluoroethylene, foam nickel (porosity: 95%∼98%, thickness: 2 mm), polyvinyl alcohol, sodium hydroxide, etc. Table 2 shows the required experimental instruments.

In Table 2, the required equipment includes an electronic balance, VDO, electrochemical workstation, constant temperature heating magnetic stirrer, vacuum filter, tablet press, slicer, and ultrasonic cleaner.

Preparation of FSCs: Before preparing FSCs, it is necessary to first prepare active materials. The preparation method for Mn-doped Ni-based active materials is as follows: First, weigh an appropriate amount of nickel chloride and manganese chloride, and prepare them into solutions of 0.6mol/L, respectively. Next, mix the nickel chloride solution and manganese chloride in a certain proportion and stir. The stirring speed is 200r/min and the stirring time is 1h. Then add 10ml of 0.2mol/L sodium carbonate solution to a mixed solution of nickel chloride and manganese chloride, and stir at room temperature for 2h. Next, transfer the obtained mixed solution to a hydrothermal reactor and react for 12h in a VDO. The temperature of the VDO is set to 50 °C. Finally, wash and filter the obtained material using anhydrous ethanol.

Preparation of FSCs EMs: First, the prepared Mn doped Ni-based active material is mixed with AcB and tetrafluoroethylene in the ratio of 80:10:10 and anhydrous ethanol, and

**TABLE 1. Materials needed for the experiment.**

| Name | Chemical formula | Purity quotient | Manufacturer |
|---|---|---|---|
| Acetylene Black (AcB) (Size 35-45 nm) | C | / | Henan Ruibo Chemical Co., LTD |
| PTFE | (C2F4)n | Analytical reagent | Chongqing New Fluorine Technology Co., LTD |
| Foam nickel (Porosity 95%~98%) | Ni | / | Kunshan Tengerhui Electronic Technology Co., LTD |
| Absolute ethyl alcohol | CH3CH2OH | 100% | Jinan Century Tongda Chemical Co., LTD |
| Nickel dichloride | NiCl2·6H2O | High-purity | Baoding Fsay Cobalt and Nickel New Material Co., LTD |
| Sodium carbonate | Na2CO3 | Analytical reagent | Tianjin Kio Chemical Reagent Company |
| Caustic potash | KOH | 85% | Tianjin Hengxing Chemical Reagent Manufacturing Co., LTD |
| Deionized water | H2O | 100% | Self prepared |
| Manganese chloride | MnCl2·4H2O | 99% | Baoding Fsay Cobalt and Nickel New Material Co., LTD |
| Sodium hydrate | NaOH | Analytical reagent | Tianjin Kio Chemical Reagent LTD |
| Polyving akohol | (CH2CHOH)n | Analytical reagent | Henan Yuheng Chemical Co., LTD |

**TABLE 2. Instruments required for the experiments.**

| Name | Type | Manufacturer |
|---|---|---|
| Electronic balance | ME104E | Beijing Zhongyi Electric Measurement Technology Co., LTD |
| Vacuum Drying Oven (VDO) | DZF-6020 | Shanghai Qixin Scientific Instrument Co., LTD |
| Electrochemical workstation | CHI760E | Shanghai Chenhua Instrument Co., Ltd |
| Thermostatic heating magnetic stirrer | DF-101S | Shijiazhuang Yangxing Instrument Trading Co., LTD |
| Vacuum filter | AP-9925N | Tianjin Otseens Instrument Co., Ltd |
| Bead machine | PC-15 | Hangzhou David Instrument Co., Ltd |
| Flaker | MSK-T10 | Kejing Zhida Technology Co., LTD |
| Ultrasonic cleaner | KQ22000E | Kunshan Ultrasound Instrument Co., LTD |

foam nickel is treated with detergent activator. The obtained active material mixture is then coated on the surface of foam nickel with a thickness of 10 mm and dried in a VDO. Then, the EM of FSCs can be obtained by slicing and pressing using a slicer and tablet press.

Preparation of FSCs: Firstly, mix the prepared EM, polyvinylidene fluoride, and AcB in an 80:10:10 ratio with anhydrous ethanol. Coat the obtained mixture onto a carbon cloth and dry it to obtain Mn-NiOOH@C flexible electrodes. Next, mix polyvinylidene fluoride, Activated Carbon (AcC), and AcB in an 80:10:10 ratio with anhydrous ethanol and coat them on a carbon cloth. After drying, the negative electrode of the capacitor can be obtained. The ratio of the positive electrode active material to the negative electrode mass of FSCs is equation (1) [10].

$$\frac{m^+}{m^-} = \frac{G^- \times \Delta V^-}{G^+ \times \Delta V^+} \tag{1}$$

In equation (1), $m^+$ and $m^-$ are the mass of AcC in the Positive and Negative (P&N) EMs. $\Delta V^+$ and $\Delta V^-$ are the voltage testing range between the P&N electrodes. $C^+$ and $C^-$ represent the SC of the three electrode test for the P&N electrodes.

Preparation of alkaline gel electrolyte: Firstly, weigh 1.5 g of polyvinyl alcohol and 15 ml of distilled water, and let it stand at room temperature for 10 minutes. Next, mix the two and heat them in a water bath while stirring at a heating temperature of 90 °C. After the polyvinyl alcohol solution appears clear, add 5ml of 6M potassium hydroxide solution and stir for one hour before cooling the solution to room temperature. Finally, the prepared P&N electrodes and alkaline gel electrolyte are assembled according to the sandwich structure to obtain FSCs.

Performance testing methods. The performance testing of the active composite electrode of FSCs includes Cyclic Voltammetry (CV) testing, Constant Current Charge Discharge (CCCD) testing, and Electrochemical Impedance Spectroscopy (EIS) testing. The CV method involves repeatedly scanning a triangular waveform at different rates by controlling the electrode potential. The potential range is to enable diverse reduction and oxidation reactions to occur alternately on the electrode, reflecting the microscopic reaction process on the electrode surface and the reversibility of electrode reactions. The SC of the tested electrode can be calculated through the CV curve, and the calculation formula is equation (2) [11].

$$\begin{cases} Q = \int_{\varphi_1}^{\varphi_2} i(\varphi)\, dt = \int_{\varphi_1}^{\varphi_2} i(\varphi)\, d\varphi / ds \\ C_s = \dfrac{Q}{2\Delta V \cdot m} \end{cases} \tag{2}$$

In equation (2), $Q$ is the total amount of charge. $\varphi_1$ and $\varphi_2$ represent the starting and ending voltages of the CV test. $i(\varphi)$ represents the current at the current voltage $\varphi$. $t$ represents time. $s$ is the scanning rate. $C_s$ is SC. $\Delta V$ represents the voltage testing interval. $m$ is the active mass. Through CCCV testing, not only can the SC of the tested electrode be calculated, but also its SC retention rate can be calculated. The formula is shown in equation (3) [12].

$$\begin{cases} \Delta C = \left| C_{low} - C_{high} \right| \\ \alpha = \dfrac{\Delta C}{C_{high}} \times 100\% \\ \beta = 1 - \alpha \end{cases} \tag{3}$$

In equation (3), $\Delta C$ and $\alpha$ represent the difference in SC and the rate of change of SC at different voltages. $C_{low}$ and $C_{high}$ represent the SC at −0.4 V and 0.4 V. $\beta$ represents the capacitance retention rate. EIS testing refers to applying small amplitude Sine Wave Disturbance Signals (SWDS) of different frequencies to an Electrode System (EcS), obtaining the electrode impedance from the correlation between the response of the EcS and the disturbance signal, and inferring the equivalent circuit of the electrode. Furthermore, the dynamic processes and mechanisms contained in the EcS can be analyzed, and the dynamic parameters of the EcS can be estimated from the parameter values of relevant components in the equivalent circuit. The connection between the excitation signal and response signal of EIS testing is equation (4).

$$Y = G(w)\,X \tag{4}$$

In equation (4), $Y$ represents the response signal. $w$ represents angular frequency. $X$ represents the excitation signal. If $G$ and $X$ are the sine wave current/voltage signal, then $Y$ represents the impedance of the system. If $X$ and $Y$ are the sine wave voltage/current signal, then $G$ represents the admittance of the system.

### B. POWER DISPATCH ASSISTED DECISION-MAKING BASED ON PPO ALGORITHM

Although the above-mentioned FSCs can achieve energy storage in the power system and reduce power waste, the randomness and volatility of RE make it difficult to predict energy output. This will lead to significant uncertainty in power dispatch, thereby increasing the power deviation of the grid. Given this, this study develops a PPO-based Power Dispatch Auxiliary Decision-Making (PDAM) method based on the introduction of FSCs. When constructing the PDAM model, to reduce the operating cost of the power grid and the consumption rate of RE, this study first constructs the minimum objective function, as shown in equation (5) [13].

$$F = \omega_{\cos t} F_{\cos t} - \omega_{con} F_{con} - \omega_{\lim} F_{\lim} \tag{5}$$

In equation (5), $F$ means that the minimum goal is achieved in the dispatch cycle when taking into account the grid operation security, RE consumption efficiency, and operation cost. $F_{\cos t}$ and $\omega_{\cos t}$ respectively represent the Unit Operating Cost (UOC) function and its weights. $F_{con}$ and $\omega_{con}$ represent the RE absorption function and its weights, respectively. $F_{\lim}$ and $\omega_{\lim}$ represent the line crossing function and its weights, respectively. The so-called consumption of RE refers to the process of effectively utilizing RE, and line exceeding the limit refers to the current in the line exceeding the rated value. The UOC function is equation (6) [14].

$$F_{\cos t,t} = \sum_{i=1}^{N} \left( a_i P_{i,t}^2 + b_i P_{i,t} + c_i \right) + C_{on,off} \tag{6}$$

In equation (6), $F_{\cos t,t}$ represents the UOC of period $t$. $N$ represents the quantity of units. $a_i$, $b_i$ and $c_i$ represent the power generation cost coefficients of the $i$-th unit. $P_{i,t}$ represents the active output of the $i$-th unit during the $t$ period. $C_{on,off}$ represents the start-up and shutdown cost of the unit. The RE absorption function is equation (7).

$$F_{con} = \sum_{i=1}^{N_{re}} P_{re,i,t} \Big/ \sum_{i=1}^{N_{re}} \bar{P}_{re,i,t} \tag{7}$$

In equation (7), $N_{re}$ represents the number of RE units. $P_{re,i,t}$ and $\bar{P}_{re,i,t}$ are the actual and maximum active outputs of the $i$-th RE unit during the $t$ period. The line crossing function is equation (8).

$$F_{\lim} = \frac{1}{N_{line}} \sum_{i=1}^{N_{line}} \min \left( \frac{I_{i,t}}{I_{i,t} + \chi}, 1 \right) \tag{8}$$

In equation (8), $F_{\lim,t}$ represents the line exceeding the limit during period $t$. $N_{line}$ represents the amount of branches in the power grid. $I_{i,t}$ represents the current flowing through $i$ branches during period $t$. $I_{i,\max}$ is the max-current allowed for $i$ branches to pass through. $\chi$ represents a constant, with a value of 0.001. For high proportion RE power grids, the constraint conditions are shown in equation (9).

$$\begin{cases} \sum_{i=1}^{N_{re}} P_{con,i,t} + \sum_{i=1}^{N_{th}} P_{th,i,t} + P_{bal,t} - \sum_{i=1}^{N_{load}} P_{load,i,t} = 0 \\ P_{th,i,\min} \le P_{th,i,t} \le P_{th,i,\max}, 0 \le P_{con,i,t} \le P_{con,i,\max} \\ 0.9 P_{bal,\min} \le P_{bal,t} \le 1.1 P_{bal,\max} \end{cases} \tag{9}$$

In equation (9), $P_{con,i,t}$ and $P_{th,i,t}$ are the active output of the $i$-th RE unit and Thermal Power Unit (TPU) during period $t$. $N_{th}$ represents the quantity of TPUs. $P_{bal,t}$ represents the active output of the balancing unit during period $t$. $N_{load}$ represents the number of loads. $P_{load,i,t}$ represents the active power consumed by the $i$-th load during the $t$ period. $P_{th,i,\min}$ and $P_{th,i,\max}$ represent the min and max active output of the $i$-th TPU. $P_{con,i,\max}$ represents the max active output of the $i$-th RE unit. $P_{bal,\min}$ and $P_{bal,\max}$ denote the min and max active power output of the balanced unit. Due to the presence of a certain number of TPUs in a high proportion of RE power grid, the output adjustment value of TPUs needs to satisfy the climbing constraint, as shown in equation (10).

$$\begin{cases} D_i \le P_{th,i,t+1} \le U_i \\ D_i = \max \left( \left( P_{th,i,\min} - P_{th,i,t} \right), rate * P_{th,i,\max} \right) \\ U_i = \min \left( \left( P_{th,i,\max} - P_{th,i,t} \right), rate * P_{th,i,\max} \right) \end{cases} \tag{10}$$

In equation (10), $D_i$ and $U_i$ are the max-downhill and max-uphill values allowed for adjustment of the $i$-th TPU. $rate$ represents the climbing rate of TPUs. After constructing the objective function and constraints of the high proportion RE power grid, the auxiliary decision-making model can be constructed. The interaction mode of the PDAM model is Figure 1.

In Figure 1, the Power Dispatch System (PDS) will determine the next action based on the current environmental state, and after executing the action, the updated state will

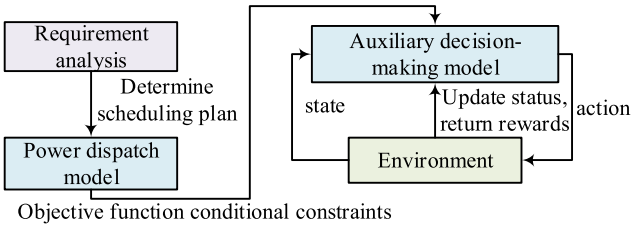



**FIGURE 1.** Square interaction mode of the PDAM model.

be returned to the system. Therefore, in PDS, state and action spaces, and return functions are the main components. Considering that the state space needs to contain dynamic changes and environmental information, which have a significant impact on algorithm performance, it is necessary to fully consider the return function when constructing the state space to ensure the reliability of generated decisions [15], [16]. The state space design is shown in Figure 2.

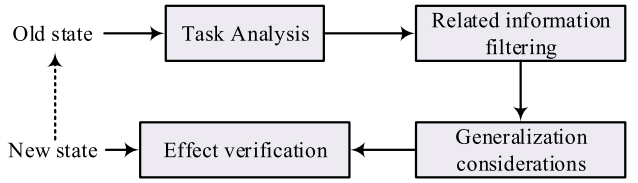



**FIGURE 2.** State space design process.

In Figure 2, the old state will obtain a large amount of relevant information after task analysis. To reduce the amount of information and computational complexity, it is necessary to filter relevant information. Then, by combining the filtered relevant information, the generalization of the algorithm is considered, and the effectiveness is verified to obtain the corresponding new state. By doing so, a reasonable state space can be constructed. In high proportion REPDS, considering the correlation between the output state of RE and the consumption reward, the constructed state space is equation (11) [17].

$$s_t = \left\{P_{th,t}, P_{con,t}, P_{bal,t}, rho_t, P_{load,t+1}, P_{con,\max,t+1}\right\} \tag{11}$$

In equation (11), $s_t$ represents the state. $P_{th,t}$ and $P_{con,t}$ represent the active output of TPUs and RE units at scheduling time $t$. $rho_t$ represents the branch power load rate. For action space, it must have the characteristics of complete timeliness and functionality, and can improve the efficiency of intelligent agents exploring the environment. Given this, the constructed action space is equation (12) [18], [19].

$$a_t = \left\{a_{th,t}, a_{con,t}, a_{bal,t}\right\} \tag{12}$$

In equation (12), $a_t$ is the action. $a_{th,t}$, $a_{con,t}$ and $a_{th,t}$ represent the active output adjustment values of TPUs, RE units, and balance units at time $t$. For the return function, it is necessary to reduce its sparsity and ensure its correspondence with the state space. In view of this, the constructed return function is equation (13) [20], [21].

$$\begin{aligned} R = {} & a_1 * r_{1,t} + a_2 * r_{2,t} + a_3 * r_{3,t} + a_4 * r_{4,t} \\ & + a_5 * r_{5,t} + a_6 * r_{6,t} \end{aligned} \tag{13}$$

In equation (13), $R$ represents the return function. $a_1$ to $a_6$ represent the dimensions of the return term. $r_{1,t}$ to $r_{6,t}$ represent the normalized value of the return term at time R. When conducting PDAM, considering the continuous action control problem of intelligent agents, this study chose the PPO algorithm as the ADM. The PPO algorithm, as a policy gradient algorithm, can handle problems in high-dimensional continuous action spaces and nonlinear policy functions and has high accuracy and stability [22], [23]. Unlike traditional policy gradient algorithms, the PPO algorithm constrains the update amplitude of policies by restricting the mechanism of policy function changes to limit the step size and ensure the stability of the optimization process. At the same time, it utilizes the Actor-Critic structure to explore and improve the environment and improve the efficiency of sample utilization [24], [25]. The proximal strategy objective optimization function of the PPO algorithm is equation (14).

$$\begin{aligned} L_{clip}(\theta) = {} & E_{(st,at)\sim\pi_\theta}\left[\min \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{old}}(a_t|s_t)} A_\pi(s,a),\right. \\ & \left. clip\left(\frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{old}}(a_t|s_t)}, 1-\varepsilon, 1+\varepsilon\right) A_\pi(s,a)\right] \\ = {} & E_{(st,at)\sim\pi_\theta}\left[\min\left(\rho_t(\theta) A_\pi(s,a)\right),\right. \\ & \left. clip\left(\rho_t(\theta), 1-\varepsilon, 1+\varepsilon\right) A_\pi(s,a)\right] \end{aligned} \tag{14}$$

In equation (14), $L_{clip}(\theta)$ represents the optimization function of the proximal strategy objective. $E_{(st,at)\sim\pi_\theta}$ represents the expected value. $\pi_\theta(a_t|s_t)$ and $\pi_{\theta_{old}}(a_t|s_t)$ are the probability of selecting action $a_t$ in state $s_t$ and after the last policy optimization. $A_\pi(s,a)$ is the advantage value of selecting action $a_t$ in state $s_t$. $\varepsilon$ represents a hyper-parameter. $\rho_t(\theta)$ represents the ratio of the probability of selecting the same action between the new and old strategies in state $s_t$. The calculation of the advantage value is equation (15) [26].

$$\begin{aligned} A_\pi(s,a) = {} & Q_\pi(s,a) - V_\pi(s) \\ = {} & E_{s'\sim P(s'|s,a)}\left[r(s) + \gamma V_\pi(s') - V_\pi(s)\right] \end{aligned} \tag{15}$$

In equation (15), $Q_\pi(s,a)$ and $V_\pi(s,a)$ are the action and state value functions. $r(s)$ represents reward. $\gamma$ is the discount factor. The network update method of PPO algorithm is shown in Figure 3.

In Figure 3, after interacting with the intelligent sports environment, the information is input into the New Strategy Network (NSN), and the actions are sampled based on the distribution of action strategies. Then, the obtained actions are input into the environment to obtain new states and rewards, and stored in the experience library. The new states are input into the evaluation network, their state and action value functions are calculated, and the network parameters are updated through the dominance function. All state combinations in the experience base are input into the new and

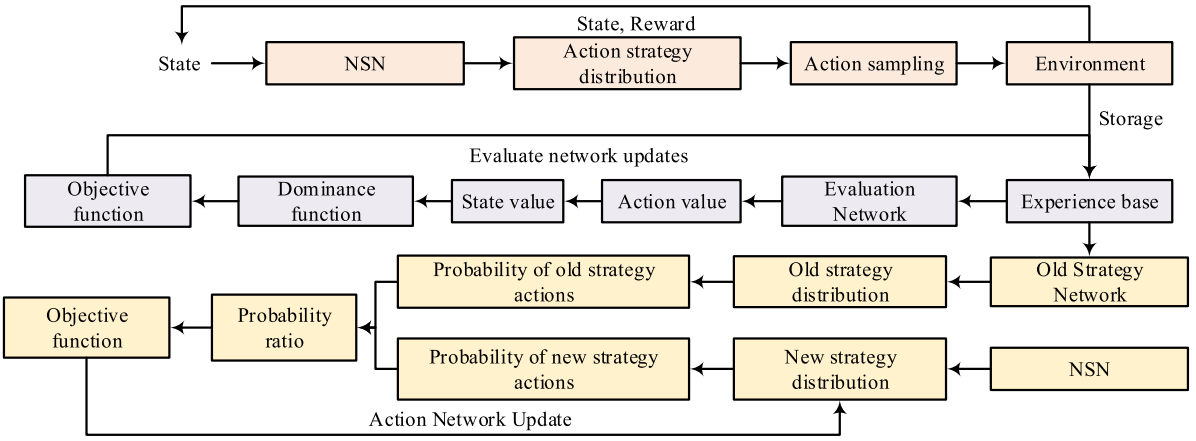



**FIGURE 3.** Network update mode of the PPO algorithm.

old policy networks, and their action probabilities are calculated based on the distribution of the new and old policies. The action probability ratio of the new and old strategies is calculated, and the NSN is updated based on the objective function. At the same time, the parameters of the NSN are used to update the old strategy network. The above operation is repeated until the algorithm converges.

## III. RESULTS

To verify the performance of high proportion RE power grid scheduling schemes based on FSCs and PPO algorithms, this study tested FSCs and PPO-based power grid scheduling ADM methods.

### A. EVALUATION OF FSC PERFORMANCE

To verify the performance of the proposed Mn&Ni-CEM-based FSCs, this study tested Mn&Ni-based Active Composite Materials (Mn&Ni-ACM) and electrodes separately. Figure 4 shows the EDS and XRD patterns of Mn&Ni-ACM.

In Figure 4 (a), Mn&Ni-ACM mainly contains Ni, Mn, O, and Cl elements, with their respective proportions of 32.9%, 3.5%, 56.0%, and 7.6%. The Ni and Mn elements have a relatively low and significant impact on the electrochemical performance of the material, while the Cl element enhances the performance. In Figure 4 (b), compared with the standard XRD spectra of NiOOH and MnO, Mn&Ni-ACM exhibit a difference of $2\theta$ MnO crystal phase appeared at both 20.2 and 21.6, and at $2\theta$ NiOOH crystal phase appeared at positions 28.2 and 37.8. Therefore, the prepared Mn&Ni-ACM is Mn-NiOOH. The CV curves of Mn-NiOOH at various reaction times and temperatures are shown in Figure 5.

In Figure 5 (a), under the condition of 50 °C, as the reaction time increases, the Peak Current of the Electrode (PCE) first decreases and then increases. When the reaction time is 12 hours, 30 hours, and 48 hours, the PCE is 0.21A, 0.08A, and 0.05A, respectively. In Figure 5 (b), when the reaction time is all 12 hours, PCE first grows and

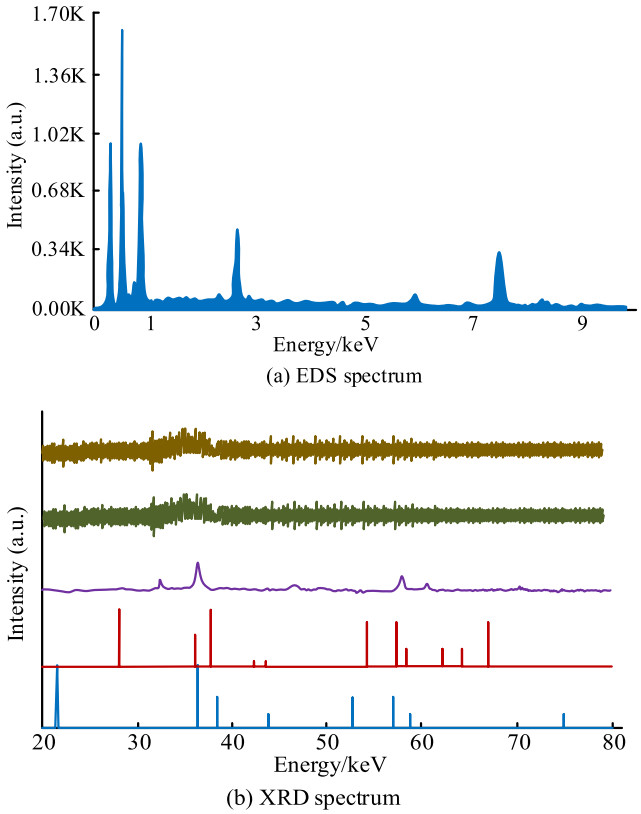
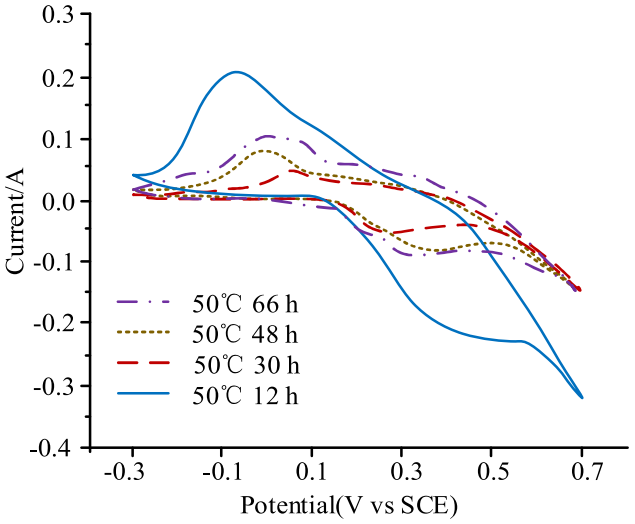
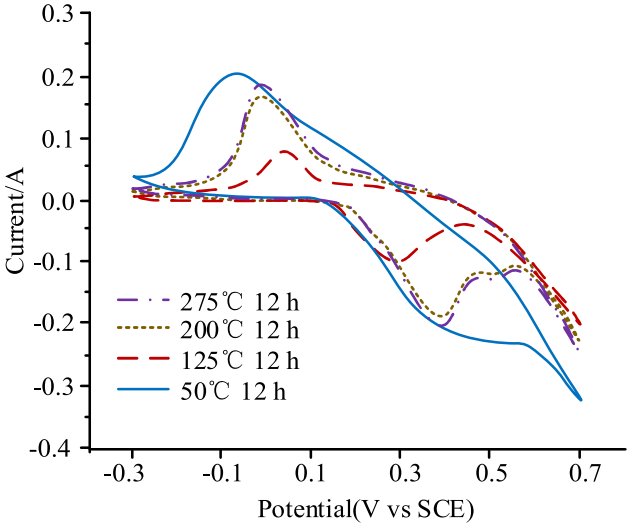



**FIGURE 4.** EDS and XRD plots of the Mn&Ni-ACM.

then lowers with the growth of reaction temperature. When the temperatures are 50 °C, 125 °C, and 200 °C, the PCE is 0.20A, 0.09A, and 0.18A, respectively. Therefore, when the reaction temperature and time are 50 °C and 12 hours, the electrochemical performance of the Mn-NiOOH is the best. In addition, Figure 5 also indicates that the Mn-NiOOH electrode mainly stores energy through redox reactions. The CCCV curves of Mn-NiOOH electrodes at different reaction times and temperatures are shown in Figure 6.

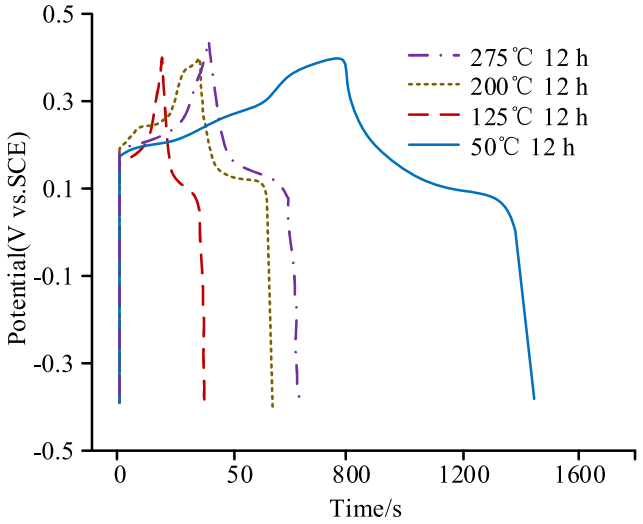



(a) CV curves at different temperatures

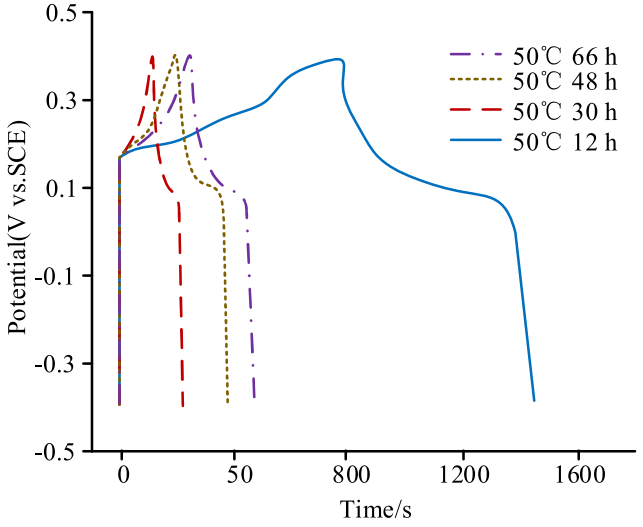



(b) CV curves at different temperatures

**FIGURE 5.** Circular voltammegram of the Mn-NiOOH electrode.




(a) CV curves at different temperatures




(b) CV curves with different reaction times

**FIGURE 6.** CCCD curves of the Mn-NiOOH electrode.

In Figure 6 (a), when the reaction time is all 12 hours, the CCCV time of the electrode decreases first and then lifts with the increase in temperature. When the reaction temperatures are 50 °C, 125 °C, and 200 °C, the charging and discharging times of the electrodes are 1450 s, 300 s, and 518 s, respectively, and the charging and discharging times are basically the same. In Figure 6 (b), when the reaction temperature is all 50 °C, the CCCV time first decreases and then increases with the increase of reaction time. When it is 12h, 30h, and 48h, the CCCV time is 1461s, 200s, and 380s, respectively. The above results indicate that the Mn-NiOOH electrode has the longest service life when the reaction temperature is 50 °C and the reaction time is 12 hours. The CCCV and SC curves under diverse CDs are shown in Figure 7.

In Figure 7 (a), the CCCV time of the electrode gradually reduces with the rise of CD. When the CD is 1A/g, the CCCV time is 1455 seconds. When the CD is 2A/g, the charging and discharging time is shortened to 521s, a reduction of 64.2%. In Figure 7 (b), as the CD rises, the SC of the electrode decreases. When the CD is 1A/g, the SC is 863F/g. When the CD is 5A/g, the SC decreases to 353F/g, a decrease of 59.1%. This indicates that the Mn-NiOOH electrode has excellent rate characteristics. The cycling performance and EIS of the Mn-NiOOH electrode are shown in Figure 8.

In Figure 8 (a), as the number of cycles increases or decreases, the SC retention of the electrode slowly decreases. When the number of cycles is 500, the SC retention rate is 80.2%. When the number of cycles is 700, the SC retention rate decreases to 75.5%. After 1000 cycles, the SC retention rate is 72.3%. In Figure 8 (b), the impedance of the Mn-NiOOH electrode is about 118 Ω, which is relatively small. This indicates that the Mn-NiOOH electrode has good electrochemical performance. This is because the bimetallic doping in the material has a synergistic effect, and foam

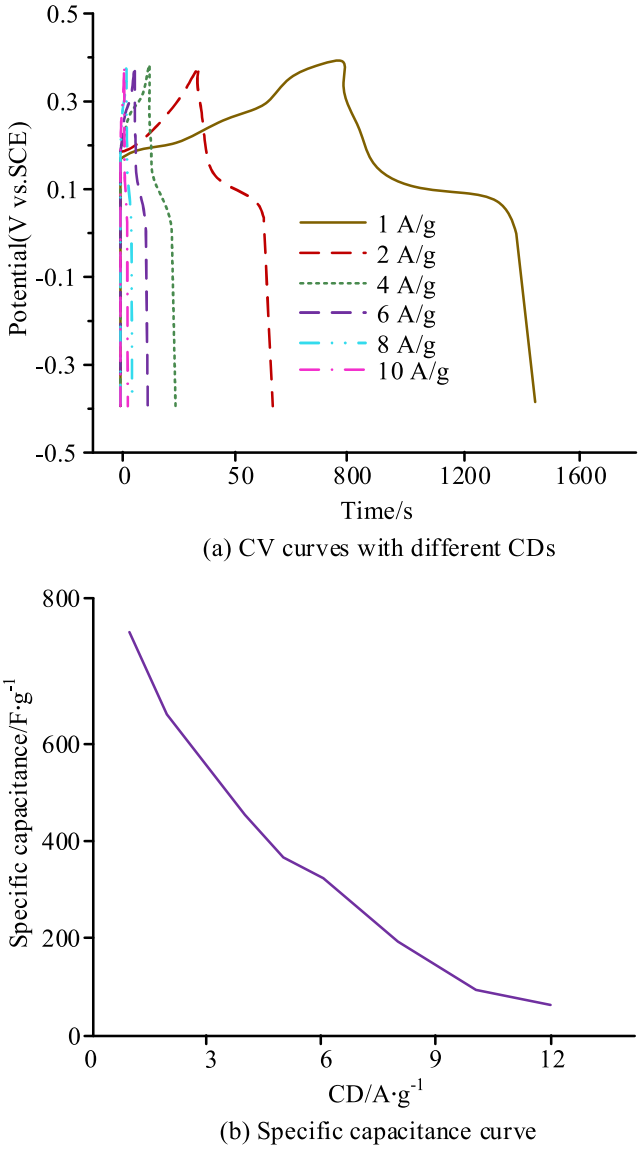



**FIGURE 7.** CCCD curves and SC curves at different CDs.

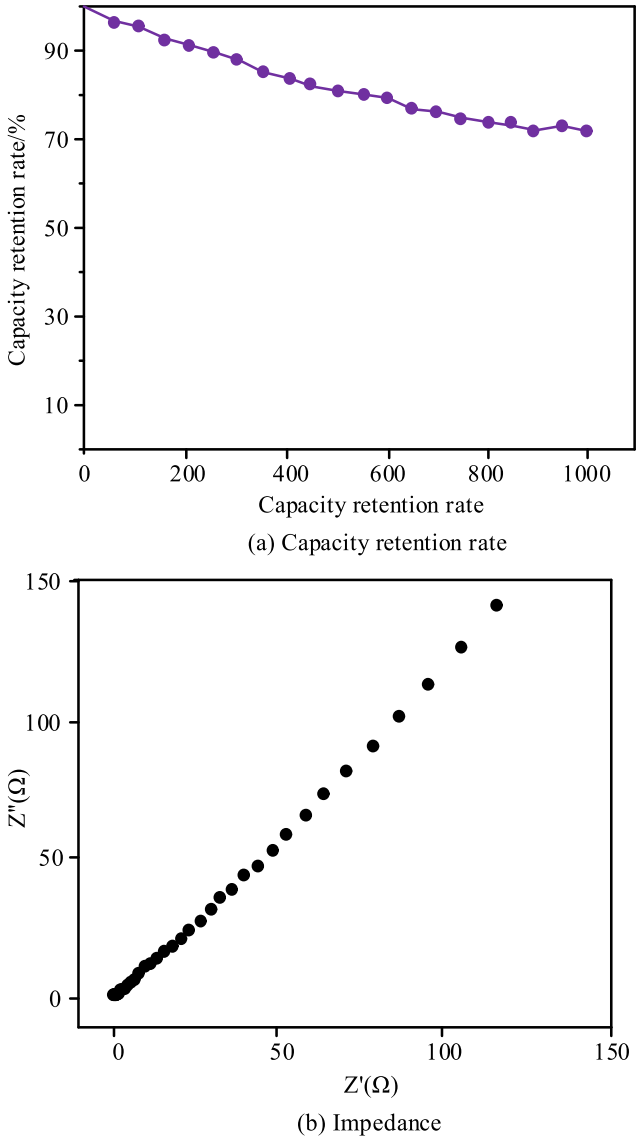



**FIGURE 8.** Circulation performance and electrochemical impedance profiles of Mn-NiOOH electrodes.

nickel provides a suitable place for electrochemical reactions. The CV and CCCV curves of FSCs are displayed in Figure 9.

In Figure 9 (a), as the scanning speed rises, the peak power of the capacitor gradually increases. When the scanning speed is 5mV/s, the peak current is 0.009 A. In Figure 9 (b), as the CD increases, the CCCV time of the capacitor gradually decreases. When the CD is 1A/g and 5A/g, the CCCV time is 128s and 10s respectively. FSCs have been proven to have good rate characteristics. The SC and ED variation curves of FSCs are exhibited in Figure 10.

In Figure 10 (a), as the CD increases, the SC of the capacitor decreases. When the CD is 1A/g, the SC is 39F/g. When the CD increases to 10A/g, the SC is 13F/g, a decrease of 66.7%. In Figure 10 (b), as the power density increases, the ED of the capacitor gradually decreases. When the power density is 8,500W/kg, the ED is 125Wh/kg. When the power density is 100,000W/kg, the ED is 41Wh/kg, a decrease of 67.2%. This indicates that FSCs have good energy storage performance.

### B. PDAM MODEL TESTING AND ANALYSIS

To ascertain the efficacy of the PDAM model in accordance with the PPO algorithm, this study conducted a series of tests. The test was conducted in a simulation environment, with 126 nodes and 54 units in the simulated power grid, including 18 RE units and 1 balancing unit. The number of loads and proficiency in load lines were 91 and 185, respectively. The data used in the experiment was sourced from the Linyi City Power Grid in Shandong Province in 2022, containing 100,000 pieces of data. In the experiment, the number of input and output neurons in the action network of the PPO algorithm and evaluation network was 348 and

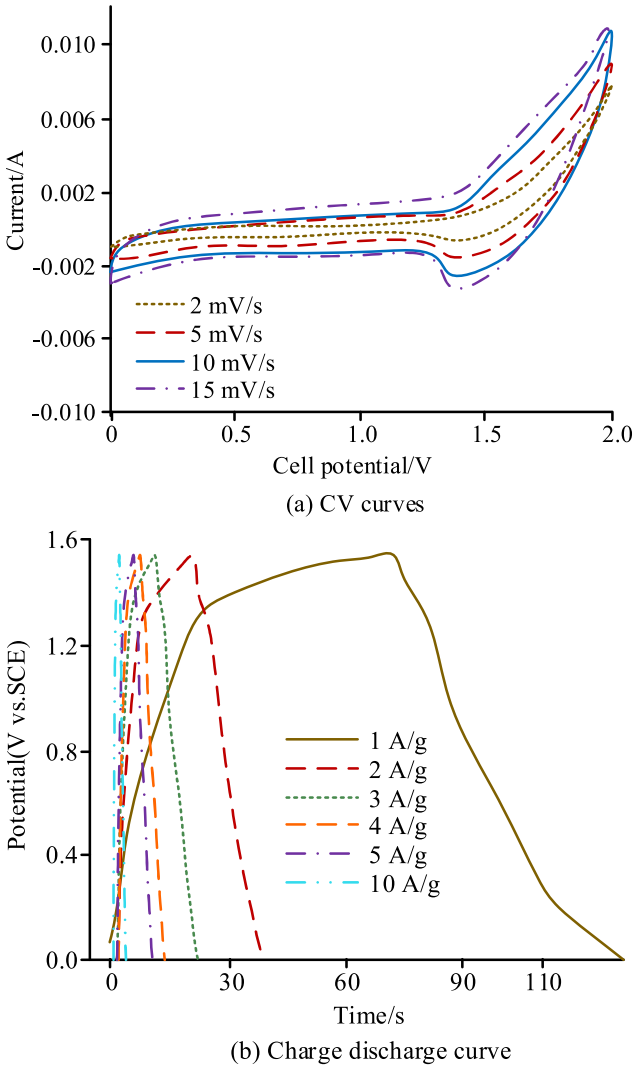



**FIGURE 9.** CV curve and CCCD curve of FSCs.

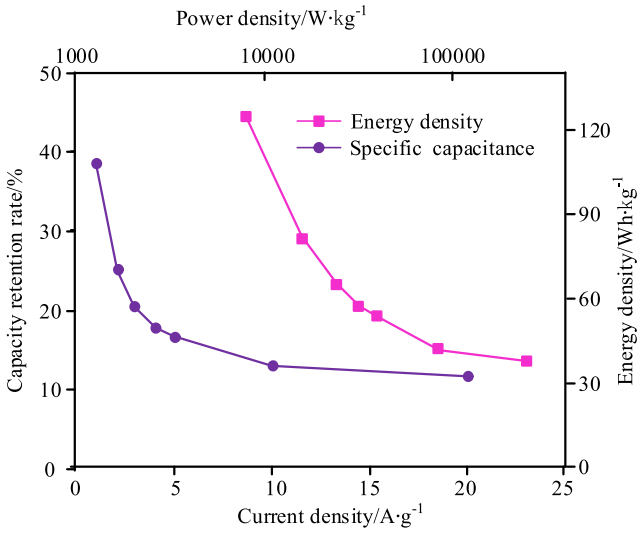



**FIGURE 10.** SC and ED change curves of FSCs.

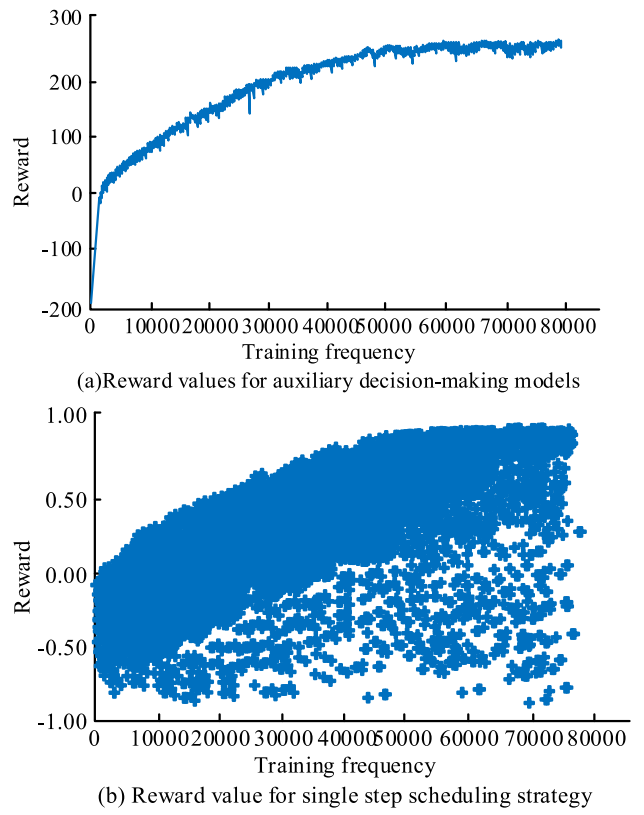



**FIGURE 11.** Reward value of the PDAM model.

108, 348 and 1, respectively. The discount factor and learning rate of the algorithm were 0.95 and 0.0001, respectively. The reward values of the PDAM model are shown in Figure 11.

In Figure 11 (a), the reward value of the PDAM model based on the PPO algorithm begins to converge after approximately 50000 iterations, at which point its reward value is approximately 257. The reward value in Figure 11 (b) increases with the number of scheduling steps. After the number of scheduling steps reaches 560000, the reward value gradually stabilizes at around 1.0. This indicates that the PDAM model based on the PPO algorithm has good ADM capability. The output situation under PPO algorithm scheduling is shown in Figure 12.

In Figure 12 (a), when the total load trend is decreasing, the output changes of TPUs, RE units, and balance units are basically consistent, all in line with the trend of active load changes in the power system, and the total output is not significantly different from the total output. In Figure 12 (b), when the total load trend is increasing, the output change trends of the three units are basically consistent. When the load tends to stabilize, due to small energy fluctuations, the output of RE units significantly increases, which is in line with the characteristics of active load changes in the power system. The PDAM model based on the PPO algorithm can formulate scheduling strategies according to the composite characteristics of the power system, reducing power differences.

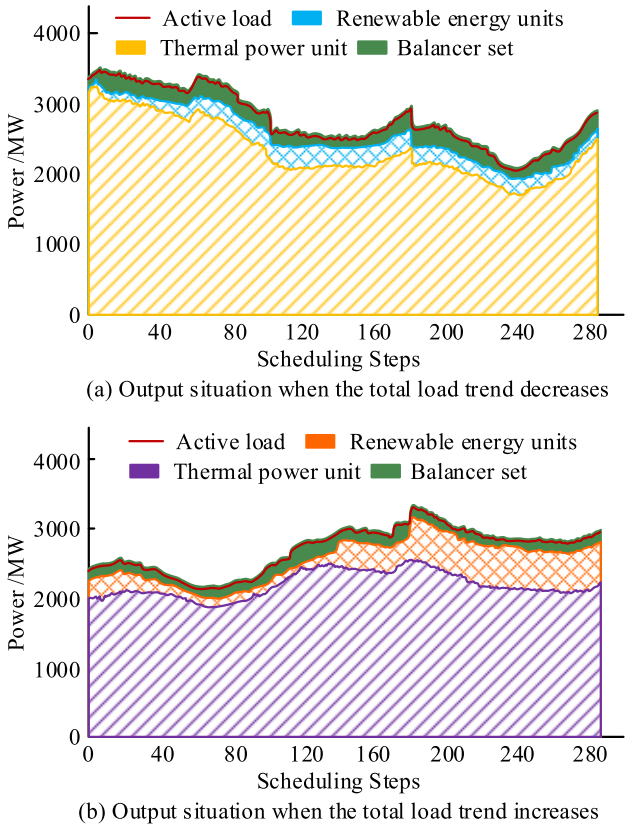



**FIGURE 12.** Output situation under PPO algorithm scheduling.

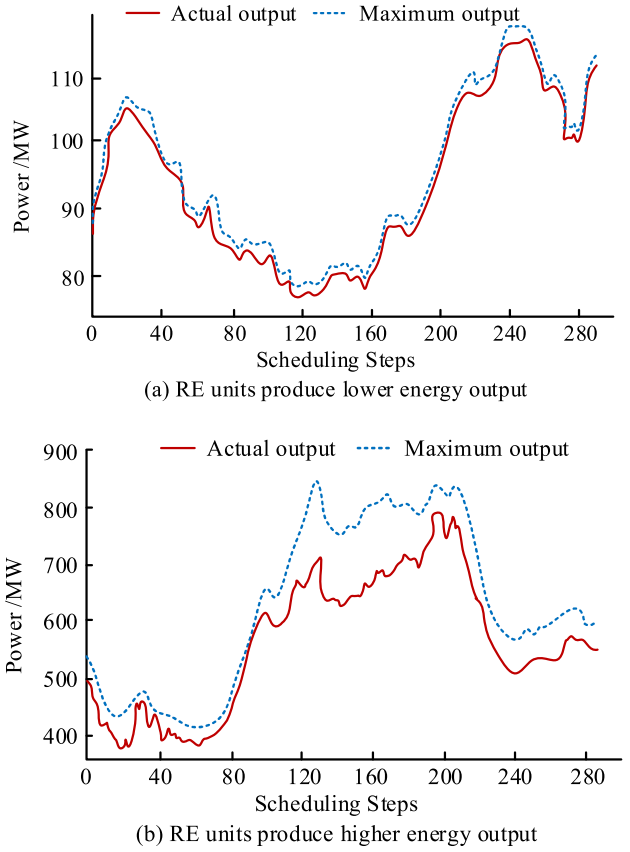



**FIGURE 13.** Consumption of renewable energy sources.

The consumption of RE under PPO algorithm scheduling is shown in Figure 13.

In Figure 13 (a), when the Energy Output Rate (EOR) of the RE unit is low, there is not much difference between its actual output and maximum output, with a fluctuation range of 75MW to 120MW, and the difference between the two does not exceed 3MW. At this time, the absorption rate of RE is relatively high, reaching 98.5%. In Figure 13 (b), when the EOR of the RE unit is high, although the trend of actual output and maximum output is consistent, there is a significant difference. When the scheduling steps are between 120-200, the difference between the actual and max output of RE significantly increases, reaching a maximum of 150MW. At this point, the absorption rate of RE is 90.1%. The comparison shows that as the EOR of RE units increases, their consumption rate gradually decreases, but overall, it still maintains a high level.

## IV. DISCUSSION

This study proposed a power dispatch method based on Mn&Ni-CEM FSCs and PPO algorithms for the scheduling problem of the RE power grid. The experimental results for FSCs based on Mn&Ni-CEM showed that the SC of Mn&Ni-CEM material was 863F/g at a CD of 1A/g, which was higher than that of flexible aluminum foil coated with carbon nanotubes [27]. At a power density of 8,500W/kg, the ED of Mn&Ni-CEM material was 125Wh/kg, significantly higher than that of FSCs based on FeCo-Ss [28]. This is because Mn&Ni-CEM materials have a nano-layered structure. Among them, the Mn-NiOOH@Ni and Mn-NiOOH@C appeared more doping levels near the Fermi level, providing more possibilities for carrier transfer and transition, and increasing the material conductivity. For the PDAM method based on the PPO algorithm, when the EOR of the RE unit was low, the actual output and maximum output were not significantly different, with a fluctuation range of 75MW to 120MW, and the difference between the two did not exceed 3MW. At this time, the absorption rate of RE was relatively high, reaching 98.5%. Compared to the comprehensive generation control and power scheduling methods based on performance-based frequency regulation markets and scheduling methods based on Coulomb and Franklin's law algorithms, it significantly reduced the power difference of the power system [29], [30]. This is because the PPO algorithm optimizes the parameters of the policy function by limiting the amplitude of policy function updates and multiple sampling, thereby achieving better performance and stability.

## V. CONCLUSION

The scheduling of high-proportion RE power systems often involves significant uncertainty, which affects the secure operation of the power grid and leads to a waste of electricity.

Given this, this study proposed a PDAM method based on FSCs and PPO algorithms to reduce power waste and improve the security of the power grid. The experiment verified that under the optimal conditions, the peak current of the EM was 0.21A, the CCCV time was the 1450s, and the maximum SC was 863F/g. The SC retention rate was 72.3% after 1,000 cycles. When the power density of FSCs with this electrode was 8,500W/kg, the ED was 125Wh/kg. Therefore, FSCs based on Mn&Ni-CEM materials had good energy storage capabilities. For the PPO algorithm, when the EOR of the RE unit was low, the actual output and maximum output were not significantly different, with a fluctuation range of 75MW to 120MW, and the difference between the two did not exceed 3MW. At this time, the absorption rate of RE was relatively high, reaching 98.5%. The above data indicates that the use of super FSCs can achieve storage of a high proportion of RE electricity, and coupled with the PPO algorithm, can well reduce the power difference of the power grid and achieve efficient energy utilization. Due to limitations in experimental conditions, only a simple and feasible hydrothermal method was used to prepare manganese nickel hybrid materials, which also has certain limitations in the packaging and testing of FSCs based on nickel manganese hybrid electrodes. Therefore, further optimization will be carried out on the packaging of FSCs.

## REFERENCES

[1] J. M. Pedraza, "The role of renewable energy in the transition to green, low-carbon power generation in Asia," *Green Low-Carbon Economy*, vol. 1, no. 2, pp. 68–84, Apr. 2023, doi: 10.47852/bonviewglce3202761.

[2] A. Deka, "The effect of forest resources, energy efficiency, and renewable energy on environmental degradation—A comparative analysis of the less- and high-emitter sub-saharan African countries," *Environ. Sci. Pollut. Res.*, vol. 30, no. 48, pp. 105781–105792, Sep. 2023, doi: 10.1007/s11356-023-29865-z.

[3] D. Sadeghi, N. Amiri, M. Marzband, A. Abusorrah, and K. Sedraoui, "Optimal sizing of hybrid renewable energy systems by considering power sharing and electric vehicles," *Int. J. Energy Res.*, vol. 46, no. 6, pp. 8288–8312, Feb. 2022, doi: 10.1002/er.7729.

[4] H. Tang, S. Wang, K. Chang, and J. Guan, "Intra-day dynamic optimal dispatch for power system based on deep Q-learning," *IEEJ Trans. Electr. Electron. Eng.*, vol. 16, no. 7, pp. 954–964, Jul. 2021, doi: 10.1002/tee.23379.

[5] Y. Huang, W. Huang, N. Tai, C. Li, R. Li, and M. Yu, "A multistage distributionally robust optimization approach for generation dispatch with demand response under endogenous and exogenous uncertainties," *IET Gener., Transmiss. Distrib.*, vol. 17, no. 22, pp. 5041–5061, Nov. 2023, doi: 10.1049/gtd2.13019.

[6] P. Vela Anandam, S. Alagarsamy, and B. Chittu Kuppusamy, "Optimal control method for power system reactive power management considering real-time uncertainties," *Optim. Control Appl. Methods*, vol. 43, no. 4, pp. 1283–1311, May 2022, doi: 10.1002/oca.2902.

[7] L. Ojeda, J. Oliva, A. I. Oliva, and C. R. Garcia, "Recycling expired pharmaceutical drugs as redox materials for efficient and sustainable flexible supercapacitors," *New J. Chem.*, vol. 47, no. 21, pp. 10090–10104, May 2023, doi: 10.1039/d3nj00497j.

[8] N. Macherla, K. Singh, M. Nerella, K. Kumari, and R. G. R. Lekkala, "Improved performance of flexible supercapacitor using naphthalene sulfonic acid-doped polyaniline/sulfur-doped reduced graphene oxide nanocomposites," *Int. J. Energy Res.*, vol. 46, no. 5, pp. 6529–6542, Jan. 2022, doi: 10.1002/er.7589.

[9] X. Wang, S. Wang, C. Li, Y. Cui, Z. Yong, D. Liang, Y. Chi, and Z. Wang, "Flexible supercapacitor based on MXene cross-linked organic gel electrolyte with wide working temperature," *Int. J. Hydrogen Energy*, vol. 48, no. 12, pp. 4921–4930, Feb. 2023, doi: 10.1016/j.ijhydene.2022.10.201.

[10] S. Zhu, Z. Zhang, J. Sheng, G. Jia, J. Ni, and Y. Li, "High-quality single-walled carbon nanotube films as current collectors for flexible supercapacitors," *J. Mater. Chem. A*, vol. 11, no. 24, pp. 12941–12949, Jun. 2023, doi: 10.1039/d2ta09396k.

[11] Y. Jiang, B. Cai, R. Xu, H. Gu, X. Qi, Z. Xu, J. Xu, and G. Liu, "One-step electrodeposition preparation of $NiCoSe_2$@carbon cloth as a flexible supercapacitor electrode material," *New J. Chem.*, vol. 47, no. 27, pp. 12649–12657, Jul. 2023, doi: 10.1039/d3nj00900a.

[12] A. M. Elshahawy, S. M. Elkatlawy, M. S. Shalaby, C. Guan, and J. Wang, "Surface-engineered $TiO_2$ for high-performance flexible supercapacitor applications," *J. Electron. Mater.*, vol. 52, no. 2, pp. 1347–1356, Feb. 2023, doi: 10.1007/s11664-022-10084-0.

[13] Z. Fu, M. Hannula, A. Jauho, K.-L. Väisänen, M. Välimäki, J. Keskinen, and M. Mäntysalo, "Cyclic bending reliability and failure mechanism of printed biodegradable flexible supercapacitor on polymer substrate," *ACS Appl. Mater. Interfaces*, vol. 14, no. 35, pp. 40145–40157, Aug. 2022, doi: 10.1021/acsami.2c08502.

[14] R. S. Karmur, D. Gogoi, M. R. Das, and N. N. Ghosh, "High-performance flexible supercapacitor device composed of a hierarchical 2-D MXene-$Ni(OH)_2$ nanocomposite and biomass-derived porous carbon electrodes," *Energy Fuels*, vol. 36, no. 15, pp. 8488–8499, Jul. 2022, doi: 10.1021/acs.energyfuels.2c01699.

[15] A. Almofleh, E. Cevik, and A. Bozkurt, "The development of novel cost-effective bio-electrolyte with glycerol host for carbon-based flexible supercapacitor applications," *Int. J. Energy Res.*, vol. 46, no. 4, pp. 5189–5199, Mar. 2022, doi: 10.1002/er.7511.

[16] Z. Pan, C. Yang, Z. Chen, and X. Ji, "Construction of $Ti_3C2T_x/WO_x$ heterostructures on carbon cloth for ultrahigh-mass loading flexible supercapacitor," *Nano Res.*, vol. 15, no. 10, pp. 8991–8999, Jul. 2022, doi: 10.1007/s12274-022-4561-6.

[17] S. Chaudhary, O. P. Sinha, and R. Mohan, "Rhinestone sheet like carbon and metal oxide-based nanocomposites for flexible supercapacitor applications," *Int. J. Energy Res.*, vol. 46, no. 7, pp. 8919–8933, Feb. 2022, doi: 10.1002/er.7767.

[18] P. Zhang and Z. Yang, "Three-dimensional Cu–Co–Se–P nanocomposites as flexible supercapacitor electrodes," *J. Mater. Science: Mater. Electron.*, vol. 33, no. 10, pp. 7396–7402, Feb. 2022, doi: 10.1007/s10854-022-07853-2.

[19] Z. Lu, H. Zhao, J. Luo, and R. Bi, "Self-supporting lignin-based carbon material flexible supercapacitor prepared by the microwave method," *Energy Fuels*, vol. 36, no. 15, pp. 8480–8487, Jul. 2022, doi: 10.1021/acs.energyfuels.2c01613.

[20] Y. Zou and H. Zhou, "Laser vision seam tracking system based on proximal policy optimization," *Ind. Robot: Int. J. Robot. Res. Appl.*, vol. 49, no. 4, pp. 770–778, Jun. 2022, doi: 10.1108/ir-08-2021-0175.

[21] R. L. Babu, S. Gurumoorthy, B. D. Parameshachari, S. C. Nelson, and Q. Hua, "End-to-end deep policy feedback-based reinforcement learning method for quantization in DNNs," *J. Circuits, Syst. Comput.*, vol. 31, no. 13, pp. 197–221, Jun. 2022, doi: 10.1142/s0218126622502322.

[22] V. Padhye, K. Lakshmanan, and A. Chaturvedi, "Proximal policy optimization based hybrid recommender systems for large scale recommendations," *Multimedia Tools Appl.*, vol. 82, no. 13, pp. 20079–20100, May 2023, doi: 10.1007/s11042-022-14231-x.

[23] Z.-C. Qiu, J.-H. Du, and X.-M. Zhang, "Vibration control of three coupled flexible beams using reinforcement learning algorithm based on proximal policy optimization," *J. Intell. Mater. Syst. Struct.*, vol. 33, no. 20, pp. 2578–2603, May 2022, doi: 10.1177/1045389x221093351.

[24] Y. Meng, S. Kuppannagari, R. Kannan, and V. Prasanna, "PPOAccel: A high-throughput acceleration framework for proximal policy optimization," *IEEE Trans. Parallel Distrib. Syst.*, vol. 33, no. 9, pp. 2066–2078, Sep. 2022, doi: 10.1109/TPDS.2021.3134709.

[25] S. Li, W. Zhang, H. Zhang, X. Zhang, and Y. Leng, "Proximal policy optimization with model-based methods," *J. Intell. Fuzzy Syst.*, vol. 42, no. 6, pp. 5399–5410, Apr. 2022, doi: 10.3233/jifs-211935.

[26] L. Capra, A. Brandonisio, and M. Lavagna, "Network architecture and action space analysis for deep reinforcement learning towards spacecraft autonomous guidance," *Adv. Space Res.*, vol. 71, no. 9, pp. 3787–3802, May 2023, doi: 10.1016/j.asr.2022.11.048.

[27] V. Ghai, K. Chatterjee, and P. K. Agnihotri, "Vertically aligned carbon nanotubes-coated aluminium foil as flexible supercapacitor electrode for high power applications," *Carbon Lett.*, vol. 31, no. 3, pp. 473–481, Jun. 2021, doi: 10.1007/s42823-020-00176-4.

[28] S.-X. Yan, S.-H. Luo, J. Feng, L. Yang, P.-W. Li, Q. Wang, Y.-H. Zhang, X. Liu, and L.-J. Chang, "Asymmetric, flexible supercapacitor based on Fe–Co alloy@sulfide with high energy and power density," *ACS Appl. Mater. Interfaces*, vol. 13, no. 42, pp. 49952–49963, Oct. 2021, doi: 10.1021/acsami.1c14537.
[29] M. Ghasemi, E. Akbari, I. Faraji Davoudkhani, A. Rahimnejad, M. B. Asadpoor, and S. A. Gadsden, "Application of Coulomb's and Franklin's laws algorithm to solve large-scale optimal reactive power dispatch problems," *Soft Comput.*, vol. 26, no. 24, pp. 13899–13923, Sep. 2022, doi: 10.1007/s00500-022-07417-w.
[30] J. Li, J. Yao, T. Yu, and X. Zhang, "Distributed deep reinforcement learning for integrated generation-control and power-dispatch of interconnected power grid with various renewable units," *IET Renew. Power Gener.*, vol. 16, no. 7, pp. 1316–1335, May 2022, doi: 10.1049/rpg2.12310.

**SHENG YANG** was born in August 1987. He received the bachelor's and master's degrees in electrical engineering from Shanghai Jiao Tong University, in 2009 and 2012, respectively. Since 2012, he has been with the Power Dispatching Control Center, Guangxi Power Grid, engaged in power dispatch operation. He is currently a Senior Engineer. He has published eight academic articles and participated in four research projects.

**JINGLONG HE** was born in February 1984. He received the bachelor's and master's degrees in electrical engineering from Hunan University, in 2004 and 2008, respectively. Since 2008, he has been with the Power Dispatching Control Center, Guangxi Power Grid. As the Manager of the Dispatch Section, engaged in power dispatch operation. He is currently a Senior Engineer. He has published 20 academic articles and participated in eight research projects.

**JINMING LIU** was born in January 1995. He received the bachelor's and master's degrees in electrical engineering from Hunan University, in 2017 and 2020, respectively. Since 2020, he has been with the Power Dispatching Control Center, Guangxi Power Grid, engaged in power dispatch operation. He is currently an Engineer. He has published two academic articles and participated in two research projects.