

Deep Reinforcement Learning-based Resilience Enhancement Framework for Distribution Networks Under Extreme Weather Events

Xiaoming Liu, *Member, IEEE*, Jun Liu, *Senior Member, IEEE, Member, CSEE*, Yu Zhao, *Student Member, IEEE, Student Member, CSEE*, and Tao Ding, *Senior Member, IEEE, Member, CSEE*

Abstract—The frequent occurrence of extreme weather events due to global warming, has led to a significant increase in operating uncertainty and significantly reduced the system's resilience. To deal with high-impact, low-probability (HILP) extreme weather events, more sophisticated control strategies and smarter methods are required to enhance the resilience of the system. In this paper, a data-driven deep reinforcement learning (DRL) framework is proposed to integrate different strategies and system situational awareness to enhance grid resilience. Specifically, the resilience enhancement problem is formulated as a Markov decision process (MDP), taking into account the situation awareness and controllability improvement of modern power systems. Next, the probabilistic effects of extreme weather events on renewable energy and transmission lines are studied and leveraged in the proposed DRL framework to improve the performance of extreme weather forecasts and estimation. Then, to speed up the training process of DRL, this paper adopts imitation learning and develops a safe topology search algorithm. Finally, an improved Soft Actor Critic (SAC) algorithm is proposed for continuous learning and training. The proposed method is tested on a modified CIGRE 15-bus medium-voltage distribution network, and the results verify the effectiveness of the proposed model and method.

Index Terms—Deep reinforcement learning, distribution networks, extreme weather events, resilience enhancement, Soft Actor-Critic.

I. INTRODUCTION

THE frequent occurrence of extreme weather events due to global warming has led to a significant increase in operating uncertainty and significantly reduced the system security [1]. Power system resilience refers to the ability of power grids to adapt to and recover quickly from disaster scenarios [2]. Therefore, it is vital to develop appropriate countermeasures to enhance grid resilience.

As an HILP event, extreme weather is considered to be one of the major threats to the security of power systems [3]. It has been observed that the grid modernization efforts in the

power industry have the potential to enhance grid resilience. In recent years, many grid modernization efforts have been made to improve the situation awareness and controllability of distribution networks. Considering the high impact of weather-related transmission permanent outages on the frequency, duration, and severity of power supply interruption and the grid modernization improvement of distribution networks, this paper focuses on the resilience enhancement of distribution networks under extreme weather events.

A great deal of research has been conducted to assess and enhance the power system resilience. In [4], the resilience response is divided into four stages, including the disaster forecast and estimation, the preventive response, the emergency response, and the resilience restoration. In [5], a novel two-stage mixed-integer optimization model is proposed to integrate preventive and emergency responses to enhance power system resilience. In [6], a novel mixed-integer linear programming model is proposed for multi-time step restoration response. Among different stages, different strategies and equipment can be used to enhance power system resilience, such as topology switching [5], generator re-dispatching or unit commitment [7], microgrids formation [8], [9], load shedding or demand response [10]. In [11], [12], an integrated optimization model is formulated to route and dispatch the repair crews and mobile power sources (MPSs) for resilience enhancement. In [13], [14], the operation and control of individual and networked microgrids are studied for grid resilience enhancement. However, the above resilience enhancement methods are all model-based, which requires precise component parameters for system modeling and some relaxed constraints for solving the optimization problem. Such limitations lead to model-based resilience enhancement methods lacking the flexibility to utilize scalable resources to enhance grid resilience, and make them unsuitable for large-scale grids.

With excellent performance in problem formulation and optimization solving, DRL-based methods have attracted the attention of scholars in resilience enhancement. In [15], a DRL-based multi-agent framework is proposed to enhance power system resilience using shunt resources. In [16], a DRL based framework is proposed to reschedule distributed generators for resilience enhancement. In [17], a DRL-based method is proposed to construct resilient distribution networks using microgrid formation. In [18], the multi-microgrid formation is used to establish a DRL-based resilience enhancement model.

Manuscript received October 31, 2022; revised January 9, 2023; accepted March 14, 2023. Date of online publication May 3, 2024; date of current version May 18, 2025. This work was supported by the National Natural Science Foundation of China (No. 52177111).

X. M. Liu, J. Liu (corresponding author, email: eeliujun@mail.xjtu.edu.cn), Y. Zhao, and T. Ding are with School of Electrical Engineering, Xi'an Jiaotong University, Xi'an 710049, China.

DOI: 10.17775/CSEEJPES.2022.07450

However, these DRL-based resilience enhancement models do not comprehensively utilize various resilience response strategies and pay little attention to the probabilistic hazard characteristics of extreme weather events.

A new DRL-based resilience enhancement framework considering the probabilistic hazard characteristics of extreme weather events is proposed in this paper. The main contributions can be summarized as follows:

- The situation awareness and controllability of modern power systems are well modeled, and a DRL-based resilience enhancement framework is proposed to integrate different strategies and system situational awareness.
- The probabilistic effects of extreme weather events on renewable energy and transmission lines are studied and leveraged in the proposed DRL framework to improve the performance of extreme weather forecasts and estimation.
- Imitation learning is adopted and a safe topology search algorithm is developed to speed up the training process of DRL.
- An improved SAC algorithm is proposed for continuous learning and training.

The rest of this paper is structured as follows. Section II briefly models the power system resilience enhancement problem. The DRL-based resilience enhancement model is established in Section III. Case studies are presented in Section IV. Section V concludes this paper.

II. PROBLEM FORMULATION

A. Power System Resilience Enhancement Framework

Power system resilience refers to the ability of power systems to adapt to disaster events and recover quickly from blackout state caused by disaster events. Based on the definition of resilience, the resilience enhancement framework under extreme weather events can be divided into five parts, as depicted in Fig. 1. The five resilience enhancement strategies are

extreme weather forecast and estimation, resilience planning, preventive response, emergency response, and restoration response, respectively. Considering that the resilience planning, which includes hardening programs, vegetation management, and deploying sensors and switches, is a long-term hardware infrastructure resilience enhancement strategy, this paper will focus on the other short-term operation and control strategies. The resilience response strategies can be defined as:

$$\begin{aligned} (EW, FP, P) &= f_{EF}(M, OS) \\ \begin{cases} (TS, GR, ESR, MPSD) = f_{PR}(FP, P, OS) & \text{if } EW = 1 \\ (TS, LS, GR) = f_{ER}(FP, P, OS) & \text{if } DW = 1 \\ (TS, GR, CLPU, MPSD) = f_{RR}(FP, P, OS) & \text{if } BW = 1 \end{cases} \end{aligned} \quad (1)$$

where EW, FP, P are the early warning status (1 for warning and 0 for security), the transmission line failure probability vector, and the predicted output vector for renewable energy; M, OS are the meteorological data and the operating state of power system; $f_{EF}(\cdot)$ represents the extreme weather forecast and estimation function; $TS, GR, ESR, MPSD, LS, CLPU$ are the topology switching, the generation re-dispatch, the energy storage re-dispatch, the mobile power source dispatch, the load shedding, and the cold load pickup, respectively; DW, BW are the disaster warning and blackout warning (1 for warning and 0 for security); f_{PR}, f_{ER}, f_{RR} are the preventive response, emergency response, and restoration response function, respectively. The specific variable definitions and function implementation are given in the subsequent sections.

B. Situation Awareness and Controllability Improvement for Modern Power Systems

To ensure the security of power systems, many grid modernization efforts have been made. The grid modernization efforts,

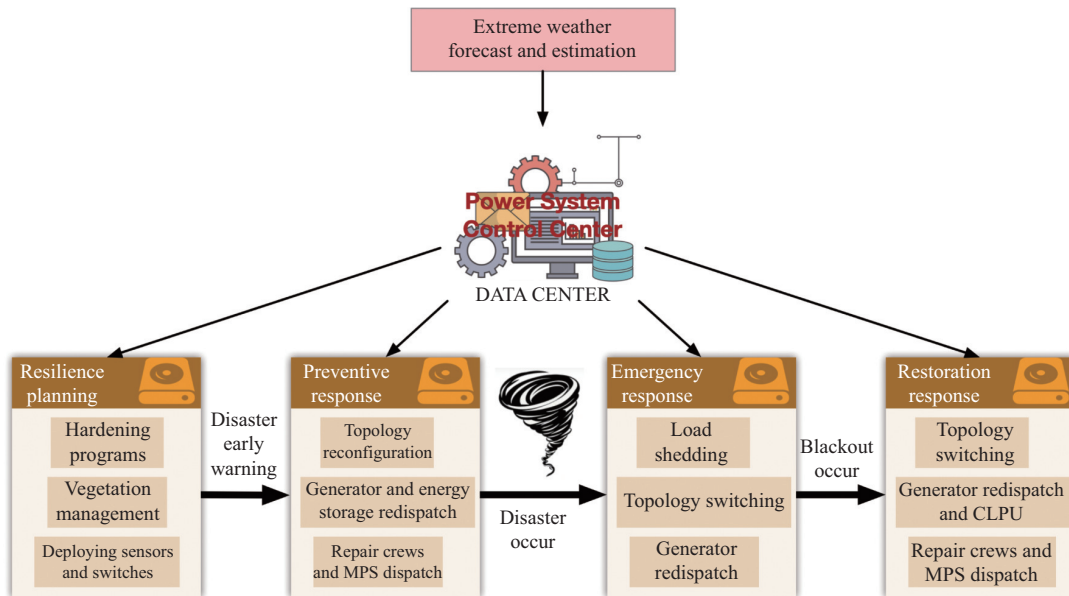


Fig. 1. Timeline of resilience enhancement framework under extreme events.

such as situation awareness and controllability improvement, provide the infrastructure to construct a smart grid and utilizing the AI technology to enhance the power system resilience.

1) Situational Awareness

The widely deployed smart meters have enabled the system operator to detect the real-time operating state of power systems. Smart meters and corresponding data acquirement systems can provide multi-source heterogeneous data such as power system data, geographical data, and weather data. Based on the situational awareness ability of power systems, the operating state OS can be described as:

$$OS = \{V, P_G, Q_G, P_L, Q_L, \text{Topo}, \rho\} \quad (2)$$

where V is the bus voltage vector, P_G is the generator active power vector, Q_G is the generator reactive power vector, P_L is the load active power vector, Q_L is the load reactive power vector, Topo is the topological representation of the power system, ρ is the loading rate vector for transmission lines and transformers. The meteorological data M contains the real-time and forecast data.

2) Grid Controllability

The development of a distribution automation system has enhanced the controllability of power systems. This paper is devoted to the comprehensive use of topology switching, load shedding, generator scheduling, energy storage scheduling, mobile power source scheduling, CLPU and other automated means to control the distribution networks and enhance the grid resilience.

In the rest of this paper, the bus index set is denoted by $B = \{b_1, b_2, \dots, b_N\}$; the mobile power source station set is denoted by $MPSS = \{MPSS_1, MPSS_2, \dots, MPSS_M\}$, where MEG_m^k is the k -th truck-mounted mobile emergency generator of $MPSS_m$; the power output at time t and maximum power output of MEG_m^k are denoted by $MP_{m,k}^t, \bar{MP}_{m,k}^t$; the distance between location i and j is denoted by $D_{i,j}$; the average velocity of MEG_m^k is denoted by v ; the current load demand for d -th load is denoted by P_L^d ; the power output and maximum power output of h -th generator are denoted by $P_G^h, \bar{P}_G^h$; the ramp rate for h -th generator is denoted by R^h .

For topology switching, in addition to switching the on/off of the transmission line, the topology of substation can also be changed by adjusting the state of isolating switches. This paper utilizes the IEEE 6-bus grid as depicted in Fig. 2(a) to explain topology of the switching. As depicted in Fig. 3, the topology adjustment from Fig. 2(a) to Fig. 2(b) can be realized by switching bus1 to bus2 of the corresponding elements in the substation $S4$.

For MPSD, two problems need to be solved, namely the routing and scheduling of MPSs. Considering that the routing problem can be solved by means of navigation software such as Google Maps, this paper focuses on the scheduling of MEGs. Assuming that any two locations are directly reachable and the distance is determined by the straight-line distance between the two locations. Then the simplified MPSD command $\{MEG_m^k, i, j\}$ can be represented as:

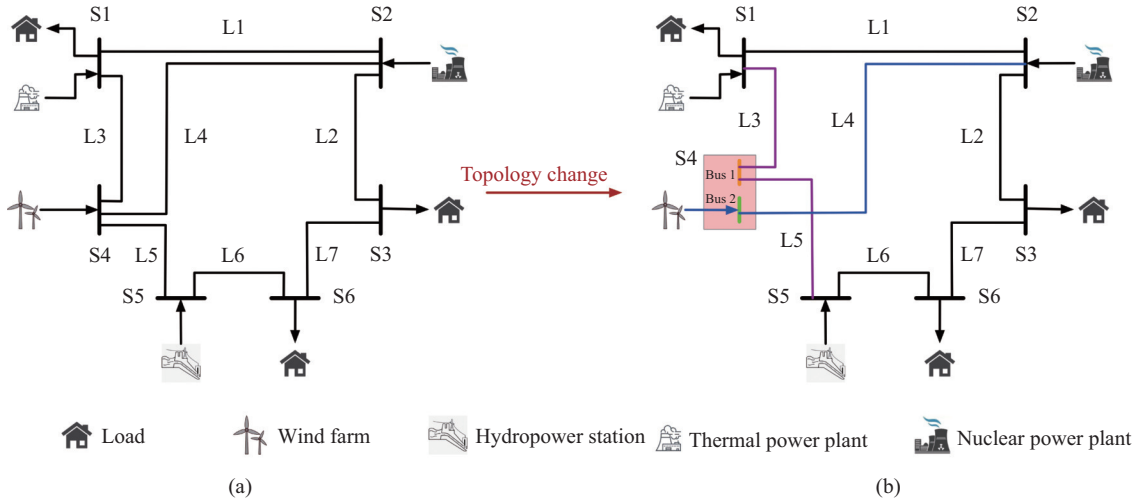


Fig. 2. Topology adjustment for IEEE 6-bus grid. (a) Original Topology. (b) Adjusted Topology.

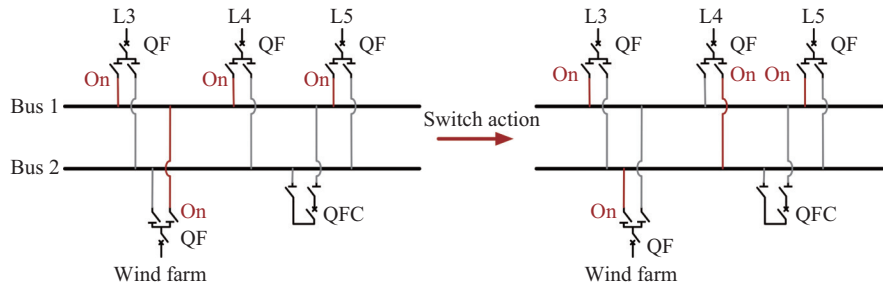


Fig. 3. Topology switching for substation $S4$.

$$\begin{cases} \text{LoM}_{m,k}^t = i \\ \text{RT} = \frac{d_{ij}}{v} \\ \text{LoM}_{m,k}^{t+\text{RT}} = j \\ i, j \in \{\text{MPSS}_m, B\}, i \neq j \end{cases} \quad (3)$$

where $\text{LoM}_{m,k}^t$ is the location of MEG_m^k at time t and RT is the run time.

For LS and GR, the constraints can be formulated as:

$$0 \leq P_{L,LS}^d \leq P_L^d, d \in D \quad (4)$$

$$P_G^h - R^h \leq P_{G,GR}^h \leq P_G^h + R^h, h \in H \quad (5)$$

$$0 \leq P_{G,GR}^h \leq \bar{P}_G^h \quad (6)$$

where $P_{L,LS}^d$ is the set value for d -th load after LS; $P_{G,GR}^h$ is the set value for h -th generator after GR.

For CLPU, the specific modeling of the load demand under CLPU conditions can refer to [3]. For ESR, the specific modeling of the energy storage system can refer to [19].

C. Modeling for Extreme Weather Events

The impact of extreme weather on the power system mainly includes two aspects: component failures and renewable energy output fluctuations. This paper mainly focuses on the weather-related line failures and weather-related generation fluctuations of wind farms.

1) Weather-Related Line Failures

The fragility function is widely used to roughly describe the component failure probability, conditioned on the loading stress s related to the potential intensity of the event. For a given event stress s , the probability of being in or exceeding a damaged state is described by the lognormal function:

$$P(s) = \Phi\left(\frac{1}{\lambda} \ln \frac{s}{\bar{S}}\right) \quad (7)$$

where $\bar{S}$ is the median value of engineering demand parameter S when the stress s reaches the threshold of the damage state; λ is the standard deviation of the natural logarithm of engineering demand parameter when the stress s reaches the threshold of the damage state; and Φ is the standard normal cumulative distribution function. The detailed parameters of fragility functions can be obtained based on expert experience and domain knowledge [20].

2) Weather-Related Output Fluctuations of Wind Farms

To study the weather-related generation fluctuations of wind farms, the wind turbine (WT) output curve [21] is introduced to quantify the relationship between wind speed and the output power:

$$P_{\text{WT}} = \begin{cases} 0, & \text{if } v < v_{\text{cut-in}} \\ \frac{1}{2} \alpha A C_p v^3, & \text{if } v_{\text{cut-in}} \leq v < v_r \\ P_r, & \text{if } v_r \leq v < v_{\text{cut-out}} \\ 0, & \text{if } v \geq v_{\text{cut-out}} \end{cases} \quad (8)$$

where P_{WT} is the output power (watts); α is the air density (kg/m^3); v is the wind speed (m/s); A is the swept area of the turbine (m^2); C_p is the power coefficient; $v_{\text{cut-in}}$, $v_{\text{cut-out}}$ are the designed cut-in and cut-out speed of the wind turbine; v_r , P_r are the rated speed and rated output power of the wind turbine.

D. MDP Formulation

With the remarkable performance in studying the sequential optimization of discrete-time stochastic systems, MDP is leveraged to model the interaction paradigm and the detailed definition can refer to [17]. In the power system resilience enhancement framework, the power system control center is treated as the agent. The operating constraints of the distribution network are considered in the environment, including power flow constraints, topology constraints, generator capacity constraints, etc. The specific definitions of state space, action space and reward function are designed as follows:

1) State Space

Based on the definition of resilience enhancement problem in (1), the state of the environment contains meteorological data, failure probability, the predicted power output for wind farm, the operating state of power system, the early warning status, the disaster warning status, and the blackout warning status:

$$S = \{M, FP, P, OS, EW, DW, BW\} \quad (9)$$

2) Action Space

Based on the definition of resilience enhancement problem in (1), the action space contains topology switching, generation dispatch, energy storage dispatch, mobile power resources dispatch, and load shedding:

$$A = \{TS, GR, ESR, MPD, LS, CLPU\} \quad (10)$$

Illustration of the improved exploration strategy.

3) Reward Function

The goal of power system resilience enhancement is to meet the load demand as much as possible under the condition that the power system operates safely and economically. The reward function is formulated as follows:

$$R = R_1(TS) + R_2(LS, GR) + R_3(ESR) + \quad (11)$$

$$R_4(MPSD) + R_5(V, \rho) \quad (12)$$

where $R_1(TS)$, $R_2(LS, GR)$, $R_3(ESR)$, $R_4(MPSD)$ are the rewards for different resilience response strategies. And $R_5(V, \rho) = \sum r_1(V_i) + \sum r_2(\rho_j)$ is the reward for evaluating the security state of the system, which can be defined as:

$$r_1(V_i) = \begin{cases} r_v, & \text{if } V_i < 0.95 \\ \frac{V_i - 0.95}{0.98 - 0.95}(\bar{r} - \underline{r}) + \underline{r}, & \text{if } 0.95 \leq V_i < 0.98 \\ \bar{r}, & \text{if } 0.98 \leq V_i < 1.02 \\ \frac{1.05 - V_i}{1.05 - 1.02}(\bar{r} - \underline{r}) + \underline{r}, & \text{if } 1.02 \leq V_i \leq 1.05 \\ r_v, & \text{if } V_i > 1.05 \end{cases} \quad (13)$$

$$r_2(\rho_j) = \begin{cases} \bar{r}, & \text{if } \rho_j < 0.8 \\ \frac{\rho_j - 0.8}{1.0 - 0.8}(\bar{r} - \underline{r}) + \underline{r}, & \text{if } 0.8 \leq \rho_j \leq 1.0 \\ r_v, & \text{if } \rho_j > 1.0 \end{cases} \quad (14)$$

where $\bar{r}$, $\underline{r}$, r_v are the maximum, minimum and violation rewards, respectively. The violation rewards will limit the voltage and loading rate within the normal range.

III. DRL-BASED RESILIENCE ENHANCEMENT MODEL

A. Framework Outline

To implement the $\{f_{PR}(\cdot), f_{ER}(\cdot), f_{RR}(\cdot)\}$, a DRL-based resilience enhancement model (DRL-REM) is proposed as shown in Fig. 4. With excellent sampling and convergence performance, SAC is adopted to train the agent. For power system operation, bad control strategies may lead to non-convergence of power flow calculations, which will greatly slow down the convergence speed of training. Thus, this paper adopts imitation learning to speed up the training process. Considering that the complex action space poses a great challenge to the training of DRL-REM, some practical tricks are introduced for action space reduction, and the specific implementation can refer to our previous work [22].

B. Imitation Learning

Imitation learning trains an agent based on a set of expert demonstrations, thereby attempting to learn an optimal policy by imitating the expert's decisions. Considering the similar problem formulation between imitation learning and DRL, the result of imitation learning can be used as an excellent initial policy for DRL. During many extreme weather events, multiple power infrastructure outages may coincide, which will pose great challenges to the power system and rapid responses are required. However, most actions in the action space contribute little to the problem and even cause non-convergence of the power flow calculation, which will greatly slow down the training process. Therefore, for extreme weather events, a parallel safe topology search algorithm is developed, as shown in Algorithm 1. In the action space, topology switching actions have a critical impact on the safe operation of power systems, so this algorithm mainly focuses on the topology switching action.

C. Soft Actor-Critic

SAC is an off-policy DRL algorithm based on the maximum-information-entropy policy architecture, which can obtain

Algorithm 1: Parallel Safe Topology Search Algorithm for Extreme Events

Initialize: parameters: $\text{Process}_n, \text{Iter}_n, \text{Topo}, \text{Action}_n, \bar{\rho}$, Security topology set Ω

```

1 for process  $\leftarrow 0$  to  $\text{Process}_n$  do
2   Action space:  $\mathbf{A} \leftarrow$  sample  $\text{Action}_n$  actions from
   original topology switching action space
3   for iter  $\leftarrow 0$  to  $\text{Iter}_n$  do
4     topo  $\leftarrow$  randomly choose a topology in  $\text{Topo}$ 
5     if  $\rho_{\max} < \bar{\rho}$  then
6       Add topo into  $\Omega$ 
7       break
8     end
9     for each action in  $\mathbf{A}$  do
10      topo' = topo  $\leftarrow$  action
11       $\rho' = \text{env} \leftarrow \text{topo}'$ 
12      if  $\rho'_{\max} < \rho_{\max}$  then
13         $\rho_{\max} = \rho'_{\max}$ 
14        topo = topo'
15      end
16    end
17  end
18  if  $\rho_{\max} < \bar{\rho}$  then
19    Add topo into  $\Omega$ 
20  else
21    Add topo into  $\text{Topo}$ 
22  end
23 end

```

stable convergence properties and high sampling efficiency [23]. The information entropy H of policy π at state s is calculated from the probabilistic distribution and can be defined as $H(\pi(\cdot|s)) = -\sum_a \pi(a|s) \ln \pi(a|s)$, where a is the action.

Unlike traditional DRL algorithms, SAC augments the conventional cumulative reward with expected information entropy, thereby introducing randomness in policy updates. The augmented objection function is defined as:

$$J(\pi) = E_{\tau \sim \pi} \left[\sum_{t=0}^T \gamma^t [(R_t + \alpha H(\pi(\cdot|s_t)))] \right] \quad (15)$$

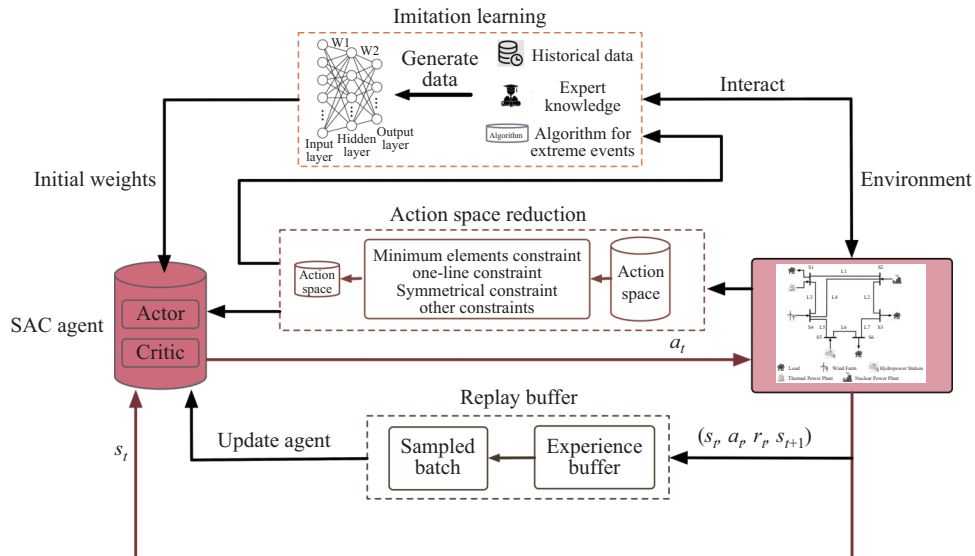


Fig. 4. DRL-based resilience enhancement framework.

where τ is the trajectory of a specific policy; γ is the discount factor; α represents the weight of information entropy, which controls the randomness of policy updates. In this way, the state value function can be calculated as:

$$V_h^\pi(s) = E_{\tau \sim \pi} \left[\sum_{t=0}^T \gamma^t (R_t + \alpha H(\pi(\cdot|s_t))) | s_0 = s \right] \quad (16)$$

and the information entropy bonuses are added to the modified action value function as:

$$Q_h^\pi(s, a) = E_{\tau \sim \pi} \left[\sum_{t=0}^T \gamma^t R_t + \alpha \sum_{t=1}^T \gamma^t H(\pi(\cdot|s_t)) | s_0 = s, a_0 = a \right] \quad (17)$$

Based on these definitions, the soft Bellman iteration can be derived as:

$$\begin{aligned} Q^\pi(s_t, a_t) &= E_{s_{t+1} \sim P} [r(s_t, a_t, s_{t+1}) + \gamma V^\pi(s_{t+1})] \\ &= E_{\substack{s_{t+1} \sim P \\ a_{t+1} \sim \pi}} [r(s_t, a_t, s_{t+1}) + \gamma (Q^\pi(s_{t+1}, a_{t+1}) \\ &\quad + \alpha H(\pi(a_{t+1}|s_{t+1})))] \\ &= r(s_t, a_t, s_{t+1}) + \alpha \gamma H(\pi(\cdot|s_{t+1})) \\ &\quad + E_{a_{t+1} \sim \pi} [\gamma Q^\pi(s_{t+1}, a_{t+1})] \\ &= r(s_t, a_t, s_{t+1}) + \gamma E_{a_{t+1} \sim \pi} [Q^\pi(s_t, a_t) \\ &\quad - \alpha \ln \pi(a_t|s_t)] \end{aligned} \quad (18)$$

where P is the distribution of the state space.

Considering that the policy distribution is parameterized and direct sampling will be non-differentiable for policy updates, a reparameterization trick is introduced to modify the implementation. As a reparameterization trick for continuous action space, the squashed Gaussian policy is defined as:

$$\tilde{a}_\theta(s, \xi) = \tanh(\mu_\theta(s) + \sigma_\theta(s) \cdot \xi), \quad \xi \sim N(0, I) \quad (19)$$

where $\xi, \mu_\theta(s), \sigma_\theta(s)$ are independent noise, mean value and standard deviation, respectively. Then, policy updates can be implemented by:

$$\max_{\theta} E_{s \sim D} [\min_{\xi \sim N} Q_{\phi_j}(s, \tilde{a}_\theta(s, \xi)) - \alpha \log \pi_\theta(\tilde{a}_\theta(s, \xi)|s)] \quad (20)$$

D. Improved Exploration Strategy and Policy Update Design

The safe operation of power grids is susceptible to major topology changes, especially under extreme events [24]. In the action space, poor topology-switching actions can lead to large-scale blackouts. During the sampling process, the original exploration strategy will inevitably select poor topology-switching actions to generate trajectories. Therefore, episodes may be terminated prematurely, which poses a significant challenge to the policy training process. In [24], an adaptive early warning mechanism is developed to help the agent choose better actions, thus enhancing the robustness of the power grid. The early warning method first samples several actions with high Q values, and then simulates and selects the optimal action. However, all sampled high-Q actions can be bad topology-switching actions, which may lead to poor convergence at the beginning of training. Thus, this paper develops an improved exploration strategy as shown in Fig. 5, which builds a reduced safe topology database for different extreme events and constructs an action discriminator (AD) to identify safe actions. Apply the action to the simulation environment, the power flow result can be obtained and the AD can be defined as follows:

$$AD = \begin{cases} \text{True,} & \text{if } \rho'_{\max} < \max\{1.0, \rho_{\max}\} \\ & \text{and } v'_{\max} < \max\{1.05, v_{\max}\} \\ & \text{and } v'_{\min} > \min\{0.95, v_{\min}\} \\ \text{False,} & \text{else} \end{cases} \quad (21)$$

where $\rho_{\max}$ is the maximum of all ρ ; $v_{\max}$ is the maximum of all V ; $\rho'_{\max}, v'_{\max}$ are the corresponding simulation results.

Applying the proposed exploration strategy to the sampling process will result in better trajectories being stored. In this way, the stability and convergence of the SAC algorithm will be greatly improved.

In addition, additional policy updates are designed for extreme events to further speed up the learning and training process. In each gradient update, samples related to extreme

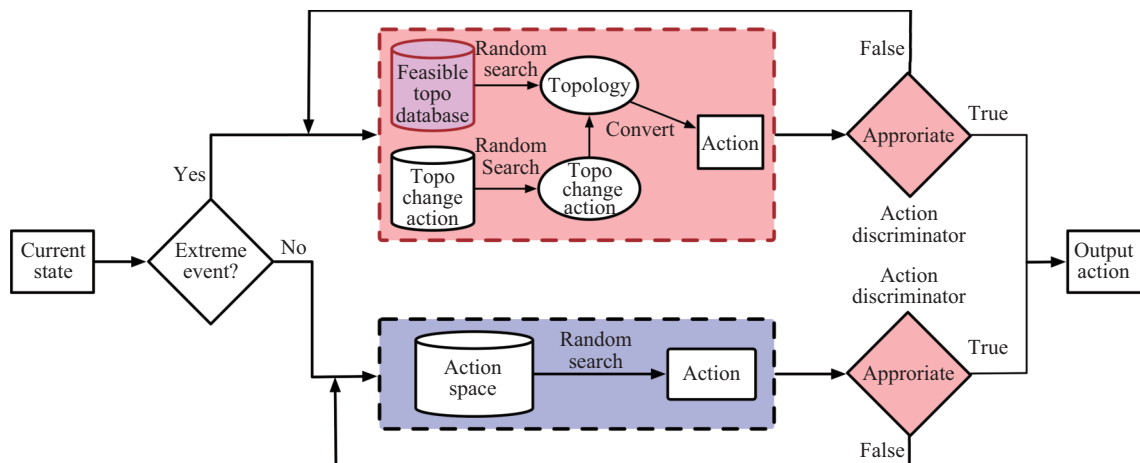


Fig. 5. Illustration of the improved exploration strategy.

events are selected, and the modified loss function is formulated as:

$$J_{\pi}(\theta) = \log \pi_{\theta}(\tilde{a}_{\theta}|s) - \log \pi_{\theta}(a_{\text{expl}}|s) \quad (22)$$

where $\tilde{a}_{\theta}$ is given by the original policy, and a_{expl} is determined by the proposed exploration strategy. Therefore, policy updates are more robust to extreme events and can be more effective in avoiding premature endings of episodes.

IV. CASE STUDIES

The proposed DRL-REM is demonstrated on a modified medium voltage distribution network from CIGRE as shown in Fig. 6. The open-source power system simulation platform, pandapower, is adopted to establish the environment. The DRL algorithms are trained on a Windows workstation with double 26-core, 2.1 GHz CPU, an RTX5000, 16 GB GPU, and 128 GB memory.

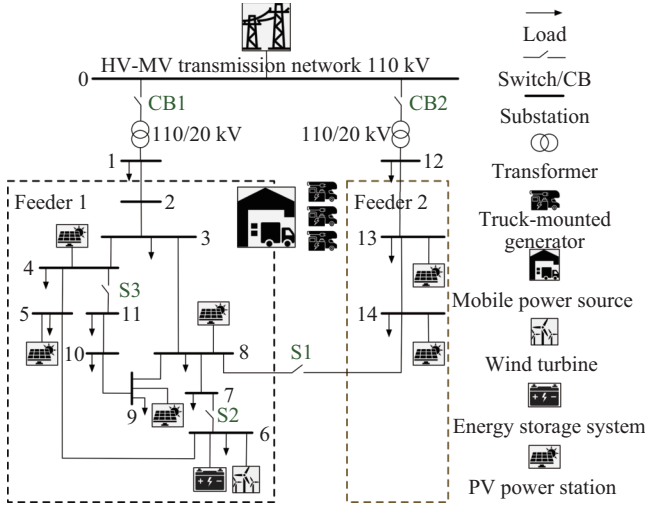


Fig. 6. One-line diagram of CIGRE MV network.

A. Description and Modification for Test System

The test network contains 15 substations, where the 4th and 5th substations are double-busbar substations. A 5 MWh ESS is added at 6-th substation and an MPSS with three 1.5 MW MEGs is deployed in this region. The average velocity of MEG is set as 60 km/h.

The test network contains 18 loads, including residential loads, commercial loads, and industrial loads. The locations and typical load demand of loads are shown in Table I, and the profiles of all types of load are shown in Fig. 7(a).

The test network contains 9 distribution energy resources (DERs) as shown in Table II, including 8 PV panels and 1 wind farm. The typical profiles of wind and solar power generations are given in Fig. 7(b).

B. Performance of Proposed Model

1) Baseline Methods

To evaluate the performance of the proposed model comprehensively, the following methods are introduced as baseline methods.

TABLE I
LOAD PARAMETERS

Type	Substation	P (MW)	Q (MVAR)
Industrial	1	14.99	3.04
Industrial	12	14.99	3.04
Residential	3	0.28	0.07
Residential	4	0.43	0.11
Residential	5	0.73	0.18
Residential	6	0.55	0.14
Residential	8	0.59	0.15
Residential	10	0.48	0.12
Residential	11	0.33	0.08
Residential	14	0.21	0.05
Commercial	1	4.85	1.59
Commercial	3	0.23	0.14
Commercial	7	0.08	0.05
Commercial	9	0.57	0.36
Commercial	10	0.07	0.04
Commercial	12	5.02	1.65
Commercial	13	0.03	0.02
Commercial	14	0.33	0.21

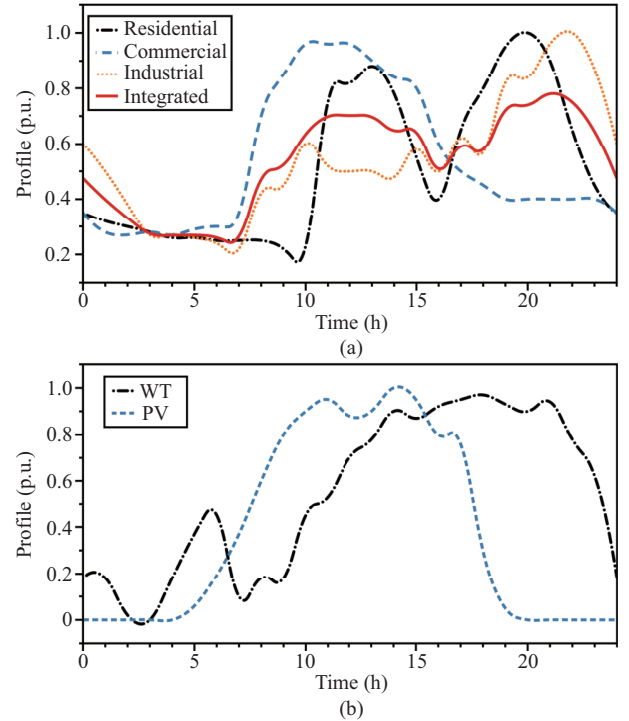


Fig. 7. Profiles for loads and DERs. (a) Load. (b) DER.

- MPC: Power system resilience enhancement under extreme weather is a sequential decision problem with uncertainty and MPC is an effective method to solve such problem. The MPC-based resilience enhancement strategy in [25] is chosen as a baseline.
- DQN: The Q-learning algorithm has been used to update the network topology for resilience enhancement of the power grid [26]. Thus, the Q-learning algorithm is chosen as a baseline to control the TS action.
- PPO: PPO, as a state-of-the-art policy gradient-based DRL method, is adopted as a baseline to control the discrete and continuous actions, including TS, GR, LS, and ESR.

TABLE II
DER PARAMETERS

Type	Substation	Capacity (MW)	Power Factor
PV	13	0.2	1
PV	4	0.2	1
PV	5	0.3	1
PV	6	0.3	1
PV	8	0.3	1
PV	9	0.3	1
PV	14	0.4	1
PV	11	0.1	1
WT	6	1.5	1

2) Training Performance Comparison of Different DRL-Based Methods

All the DRL algorithms are implemented based on the open-source framework RLib [27]. The trial-and-error strategy determines the hyperparameters of different DRL-based methods. Based on the reduced topology space and the MPC-based method, a resilience enhancement strategy can be obtained and adopted as the initial policy to pre-train the proposed DRL-REM. And the detailed training performance comparison is depicted in Fig. 8. It can be found that DQN, PPO, and DRL-REM converge after about 320, 500, 750 iterations, respectively. The average rewards converge at 1090, 1120, and 1520, respectively. Compared to DQN, the PPO-based method has a slight increase in average reward, implying that TS is the main strategy for enhancing grid resilience. By means of imitation learning, the proposed DRL-REM has a good average reward at the beginning of training and has good convergence performance. And by introducing MPSD, the proposed DRL-REM has a great improvement in the optimal average reward.

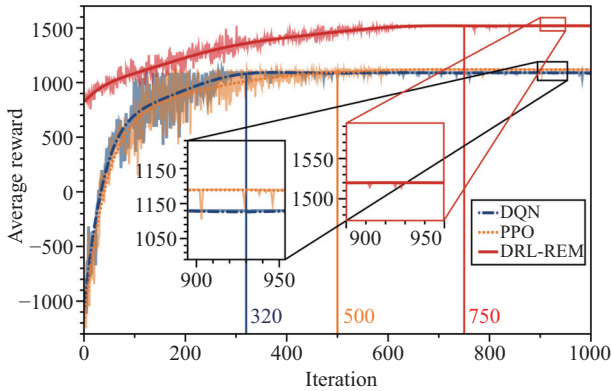


Fig. 8. Training performance comparison of different DRL-based methods.

C. Resilience Enhancement Under Typhoon Condition

To further validate the resilience enhancement performance under extreme events, a typhoon scenario is constructed as depicted in Fig. 9. The wind speed and failure probability curves of affected lines are shown in Fig. 10. The simulation interval is set to 10 minutes, and broken lines will be repaired after 2.5 hours.

All the well-trained baseline methods and our proposed DRL-REM are used to control the power grid, and the performance comparison is shown in Fig. 11. It can be seen that

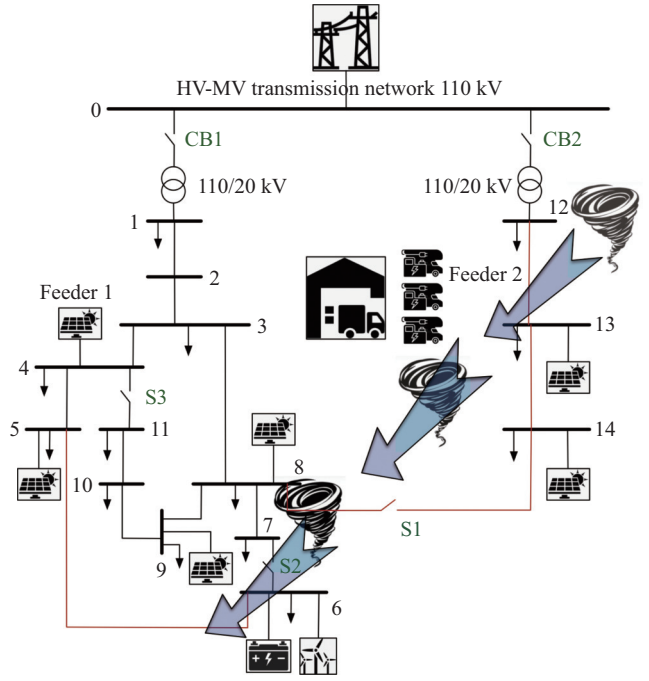


Fig. 9. Illustration of the typhoon path.

the MPC-based method mainly uses the emergency response strategy to control the grid. And all DRL-based methods can perform preventive response strategy for better performance. For system modeling, MPC-based methods are model-based and require precise component parameters. Therefore, predicted component failure probabilities and mobile power source dispatch (MPSD) are not modeled and applied to MPC-based methods. For computational efficiency, the MPC-based method and the proposed DRL-REM take an average of 300 s and 0.001 s, respectively, to provide the control policy. Compared to DQN, the PPO-based method can further schedule storage and achieve better performance when some DERs are disconnected. For the proposed DRL-REM, the introduction of MPSD further enhances grid resilience, especially when line outages would lead to the appearance of islands.

The scheduling sequence optimization is of great significance to the power grid control. Therefore, the detailed sequential control actions of our proposed DRL-REM and the corresponding power grid state trajectories between 18:20 and 23:40 are shown in Figs. 12 and 13. From the sequential TS actions in Fig. 12(a), it can be seen that our proposed model turns on the S_1 and S_2 as preventive responses. In this way, the two feeders are connected to each other, and the distribution network operates in a ring manner, avoiding sudden local power outages due to line interruptions. From the sequential MPSD actions in Fig. 12(a), it can be seen that our proposed model dispatched an MEG to bus 13 three steps ahead. Then bus 13 and bus 14 will operate in island mode during the outages of *line 8–14* and *line 12–13*. As can be seen from Fig. 12(a), the WT would be disconnected when the wind speed reaches $v_{\text{cut-out}}$. From the sequential ESR actions in Fig. 12(a), it can be seen that our proposed model schedules ESS to be charged when extreme events are imminent and to discharge during power outages. As can be seen from Fig. 13,

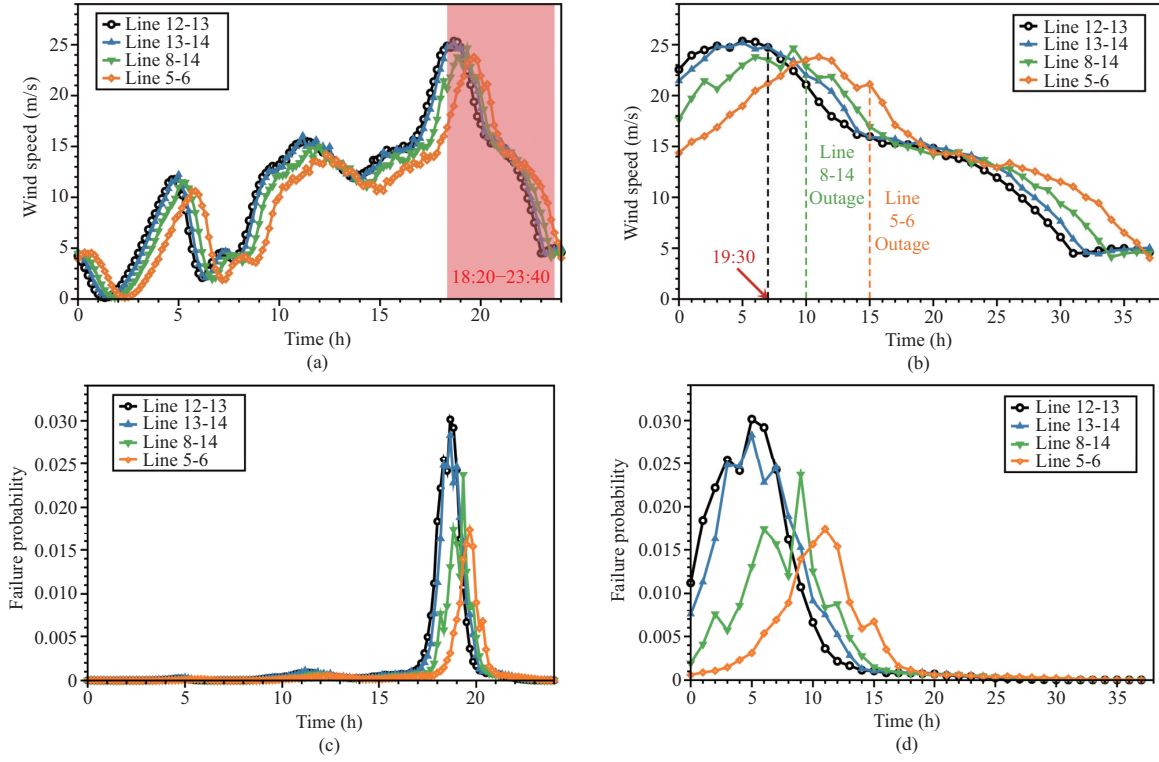


Fig. 10. Wind speed and failure probability profiles for affected lines. (a) Wind speed of all day. (b) Wind speed during the study period. (c) Line failure probability of all day. (d) Line failure probability during the study period.

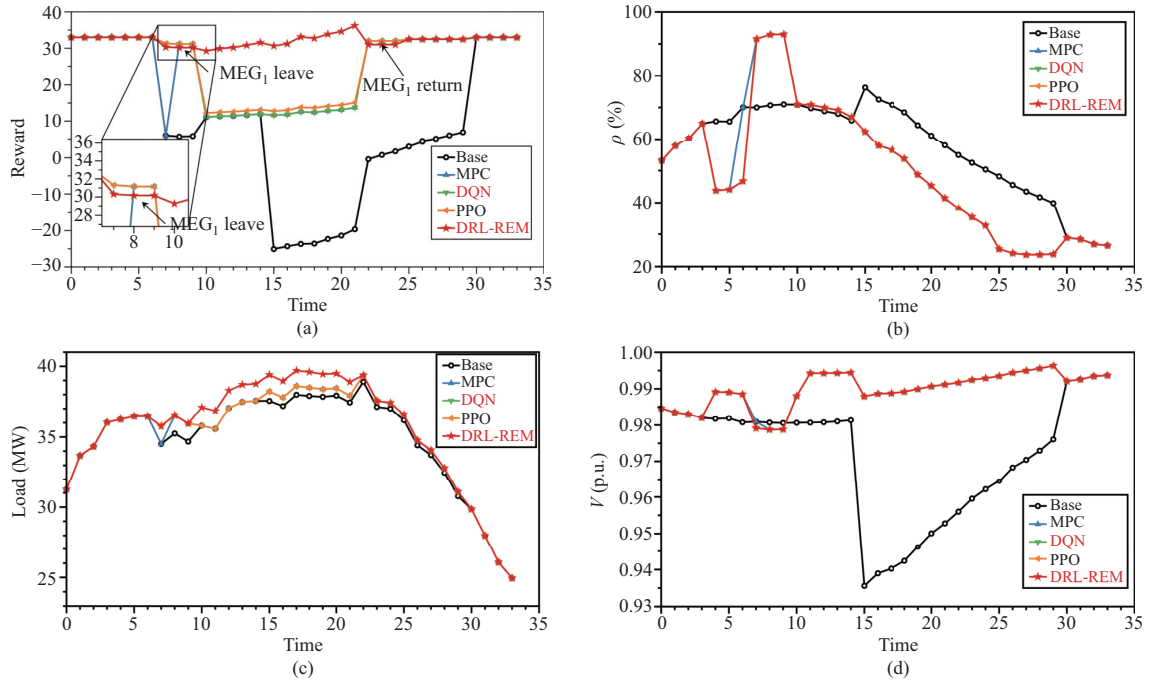


Fig. 11. Resilience performance comparison of different methods. (a) Reward. (b) Maximum ρ . (c) Load served. (d) Minimum bus voltage.

our proposed model can guarantee the operation of the grid within a safe range under extreme weather events.

V. CONCLUSION

In this paper, a data-driven DRL framework is proposed to integrate different strategies and system situational awareness

to enhance grid resilience. The probabilistic effects of extreme weather events on renewable energy and transmission lines are studied and exploited to improve the extreme weather forecast and estimation performance of the proposed model. An improved SAC algorithm is developed to accelerate the training process. The case study results show that, compared with

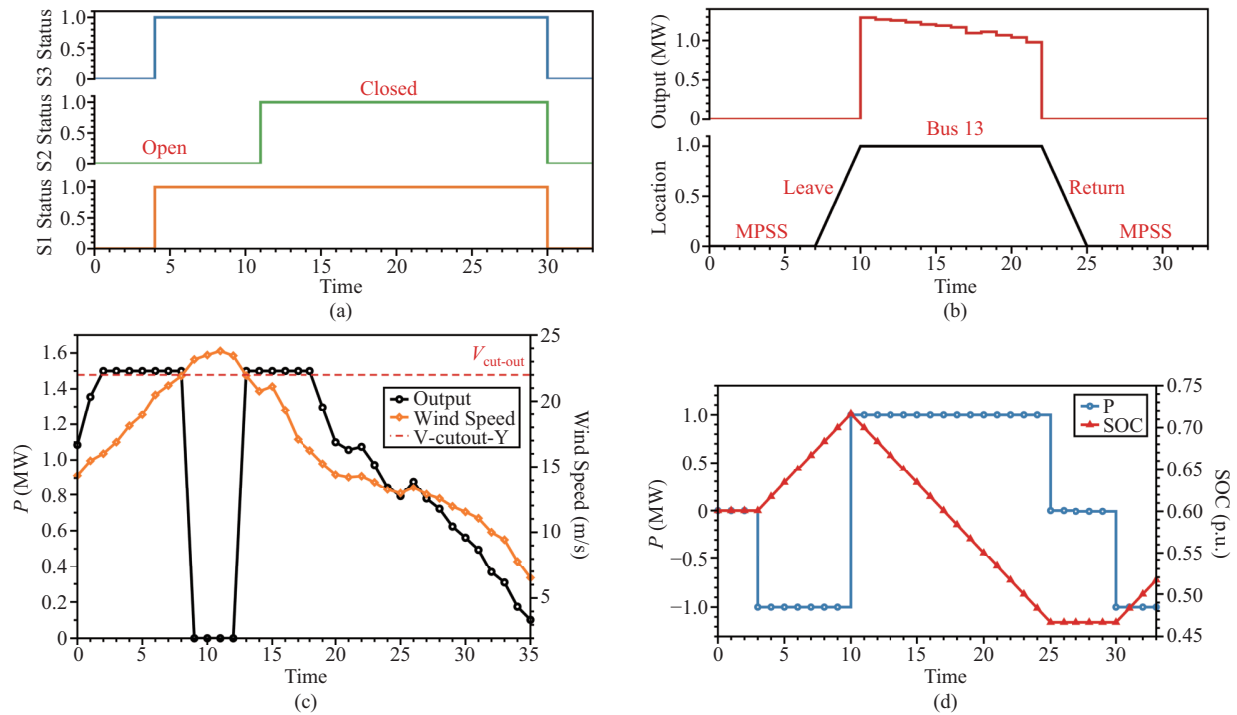


Fig. 12. The sequential control actions of DRL-REM. (a) Switch. (b) MPS. (c) Wind farm Output. (d) Storage.

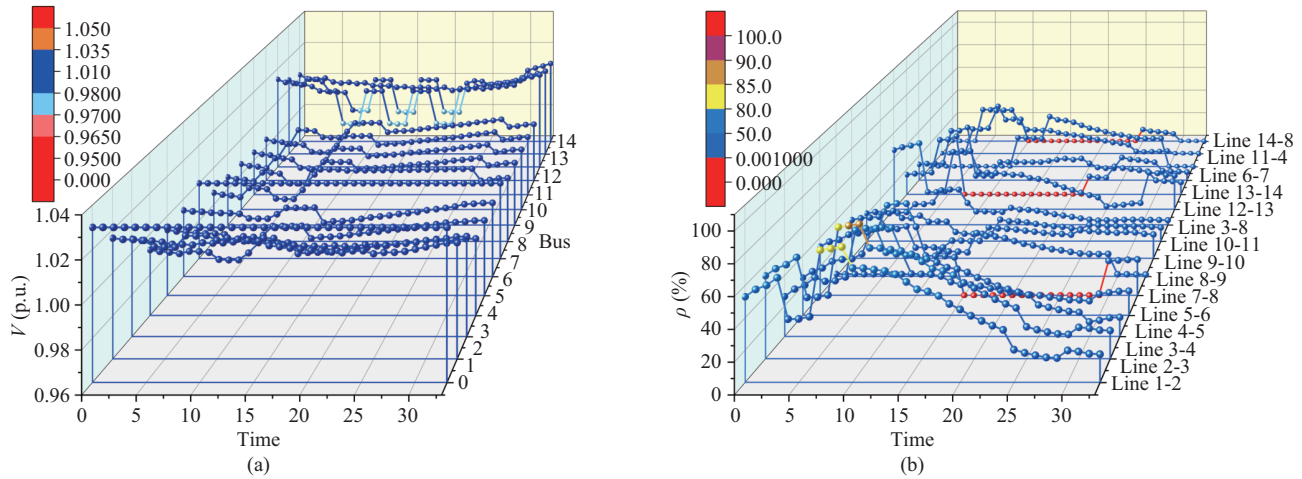


Fig. 13. The power grid state trajectory. (a) Bus voltage. (b) Line loading rate ρ .

existing methods, the proposed model can integrate different strategies and give resilience responses in advance to enhance grid resilience. It is also found that before the occurrence of extreme weather events, the proposed DRL-REM can operate the distribution network in a weakly meshed manner to avoid sudden local power outages due to line interruptions.

This paper validates the feasibility and efficiency of the DRL-based method in leveraging the situation awareness and controllability of distribution networks for resilience enhancement. It also opens new directions for developing smarter resilience controllers.

REFERENCES

- [1] Y. Yuan, H. Zhang, H. Z. Cheng, and Z. Wang, "Resilience-oriented transmission expansion planning with optimal transmission switching under typhoon weather," *CSEE Journal of Power and Energy Systems*, vol. 10, no. 1, pp. 129–138, Jan. 2024.
- [2] U.S. Department of Energy. (2010, Aug.). Hardening and Resiliency: U.S. Energy Industry Response to Recent Hurricane Seasons. [Online]. Available: <http://www.oe.netl.doe.gov/docs/HR-Report-final-081710.pdf>
- [3] C. Chen, J. H. Wang, and D. Ton, "Modernizing distribution system restoration to achieve grid resiliency against extreme weather events: An integrated solution," *Proceedings of the IEEE*, vol. 105, no. 7, pp. 1267–1288, Jul. 2017.
- [4] Y. Z. Wang, C. Chen, J. H. Wang, and R. Baldick, "Research on resilience of power systems under natural disasters—A review," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1604–1613, Mar. 2016.
- [5] G. Huang, J. H. Wang, C. Chen, J. J. Qi, and C. X. Guo, "Integration of preventive and emergency responses for power grid resilience enhancement," *IEEE Transactions on Power Systems*, vol. 32, no. 6, pp. 4451–4463, Nov. 2017.

- [6] B. Chen, C. Chen, J. H. Wang, and K. L. Butler-Purpy, "Multi-time step service restoration for advanced distribution systems and microgrids," *IEEE Transactions on Smart Grid*, vol. 9, no. 6, pp. 6793–6805, Nov. 2018.
- [7] T. Ding, M. Qu, Z. K. Wang, B. Chen, C. Chen, and M. Shahidehpour, "Power system resilience enhancement in typhoons using a three-stage day-ahead unit commitment," *IEEE Transactions on Smart Grid*, vol. 12, no. 3, pp. 2153–2164, May 2021.
- [8] C. Chen, J. H. Wang, F. Qiu, and D. B. Zhao, "Resilient distribution system by microgrids formation after natural disasters," *IEEE Transactions on Smart Grid*, vol. 7, no. 2, pp. 958–966, Mar. 2016.
- [9] S. B. Lei, C. Chen, Y. Song, and Y. H. Hou, "Radiality constraints for resilient reconfiguration of distribution systems: Formulation and application to microgrid formation," *IEEE Transactions on Smart Grid*, vol. 11, no. 5, pp. 3944–3956, Sep. 2020.
- [10] H. Masrur, M. M. Gamil, M. R. Islam, K. M. Muttaqi, M. S. H. Lipu, and T. Senjyu, "An optimized and outage-resilient energy management framework for multicarrier energy microgrids integrating demand response," *IEEE Transactions on Industry Applications*, vol. 58, no. 3, pp. 4171–4180, May/Jun. 2022.
- [11] S. B. Lei, C. Chen, Y. P. Li, and Y. H. Hou, "Resilient disaster recovery logistics of distribution systems: Co-optimize service restoration with repair crew and mobile power source dispatch," *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 6187–6202, Nov. 2019.
- [12] Z. G. Ye, C. Chen, B. Chen, and K. Wu, "Resilient service restoration for unbalanced distribution systems with distributed energy resources by leveraging mobile generators," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 2, pp. 1386–1396, Feb. 2021.
- [13] B. Chen, J. H. Wang, X. N. Lu, C. Chen, and S. J. Zhao, "Networked microgrids for grid resilience, robustness, and efficiency: A review," *IEEE Transactions on Smart Grid*, vol. 12, no. 1, pp. 18–32, Jan. 2021.
- [14] A. Shaker, A. Safari, and M. Shahidehpour, "Reactive power management for networked microgrid resilience in extreme conditions," *IEEE Transactions on Smart Grid*, vol. 12, no. 5, pp. 3940–3953, Sep. 2021.
- [15] M. Kamruzzaman, J. J. Duan, D. Shi, and M. Benidris, "A deep reinforcement learning-based multi-agent framework to enhance power system resilience using shunt resources," *IEEE Transactions on Power Systems*, vol. 36, no. 6, pp. 5525–5536, Nov. 2021.
- [16] Z.-C. Zhou, Z. Wu, and T. Jin, "Deep reinforcement learning framework for resilience enhancement of distribution systems under extreme weather events," *International Journal of Electrical Power & Energy Systems*, vol. 128, pp. 106676, Jun. 2021.
- [17] Y. X. Huang, G. F. Li, C. Chen, Y. H. Bian, T. Qian, and Z. H. Bie, "Resilient distribution networks by microgrid formation using deep reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 13, no. 6, pp. 4918–4930, Nov. 2022.
- [18] J. Zhao, F. X. Li, S. Mukherjee, and C. Sticht, "Deep reinforcement learning-based model-free on-line dynamic multi-microgrid formation to enhance resilience," *IEEE Transactions on Smart Grid*, vol. 13, no. 4, pp. 2557–2567, Jul. 2022.
- [19] Y. S. Luo, C. Chen, R. Kadavil, B. Liaw, E. Muljadi, X. F. Wu, S. K. Srivastava, T. Mosier, and E. Dufek, "A novel framework for optimizing ramping capability of hybrid energy storage systems," *IEEE Transactions on Smart Grid*, vol. 12, no. 2, pp. 1651–1662, Mar. 2021.
- [20] M. Panteli, C. Pickering, S. Wilkinson, R. Dawson, and P. Mancarella, "Power system resilience to extreme weather: Fragility modeling, probabilistic impact assessment, and adaptation measures," *IEEE Transactions on Power Systems*, vol. 32, no. 5, pp. 3747–3757, Sep. 2017.
- [21] S. Sulaeman, M. Benidris, J. Mitra, and C. Singh, "A wind farm reliability model considering both wind variability and turbine forced outages," *IEEE Transactions on Sustainable Energy*, vol. 8, no. 2, pp. 629–637, Apr. 2017.
- [22] X. M. Liu, J. Liu, Y. Zhao, and J. C. Liu, "A deep reinforcement learning framework for automatic operation control of power system considering extreme weather events," in *2022 IEEE Power & Energy Society General Meeting*, 2022, pp. 1–5.
- [23] T. Haarnoja, H. Tang, P. Abbeel, and S. Levine, "Reinforcement learning with deep energy-based policies," in *Proceedings of the 34th International Conference on Machine Learning*, 2017, pp. 1352–1361.
- [24] T. Lan, J. J. Duan, B. Zhang, D. Shi, Z. W. Wang, R. S. Diao, and X. H. Zhang, "AI-based autonomous line flow control via topology adjustment for maximizing time-series ATCs," in *2020 IEEE Power & Energy Society General Meeting*, 2020, pp. 1–5.
- [25] A. T. Eseye, X. Y. Zhang, B. Kneven, and W. Jones, "Enhancing distribution grid resilience through model predictive controller enabled prioritized load restoration strategy," in *2020 52nd North American Power Symposium (NAPS)*, 2021, pp. 1–6.
- [26] Q. M. Li, X. Zhang, J. B. Guo, X. W. Shan, Z. W. Wang, Z. Li, and C. K. Tse, "Integrating reinforcement learning and optimal power dispatch to enhance power grid resilience," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 69, no. 3, pp. 1402–1406, Mar. 2022.
- [27] E. Liang, R. Liaw, R. Nishihara, P. Moritz, R. Fox, K. Goldberg, J. Gonzalez, M. Jordan, and I. Stoica, "RLlib: Abstractions for distributed reinforcement learning," in *Proceedings of the 35th International Conference on Machine Learning*, 2018, pp. 3053–3062.



Xiaoming Liu received his B.S. and Ph.D. degrees all in Electrical Engineering from Xi'an Jiaotong University, Xi'an, China, in 2019 and 2025, respectively. Currently, he is an Assistant Professor at School of Electrical Engineering, Dalian University of Technology. His research interests mainly include power system risk assessment, power system operation and control, Bayesian deep learning, smart grids, machine learning and AI applications in power and energy systems.



control, smart grids, and integrated energy systems.

Jun Liu received his B.S. and Ph.D. degrees all in Electrical Engineering from Xi'an Jiaotong University, Xi'an, China, in 2004 and 2012, respectively. Currently, he is a Professor at the School of Electrical Engineering, Xi'an Jiaotong University. From October 2008 to August 2010, he was a visiting scholar at Department of Electrical and Computer Engineering in Texas A&M University, College Station, Texas, US. His research interests include power system stability, renewable energy and EV integration, power system operation and



Yu Zhao received the B.S. degree in Electrical Engineering from Xi'an Jiaotong University, Xi'an, China, in 2021. He is currently working towards the Ph.D. degree in Electrical Engineering at Xi'an Jiaotong University, Xi'an, China. His research interests mainly include power system operation and control, cyber-physical systems, and artificial intelligence applications in energy/power systems.



Tao Ding received the B.S.E.E. and M.S.E.E. degrees from Southeast University, Nanjing, China, in 2009 and 2012, respectively, and the Ph.D. degree from Tsinghua University, Beijing, China, in 2015. From 2013 to 2014, he was a Visiting Scholar with the Department of Electrical Engineering and Computer Science, University of Tennessee, Knoxville, TN, USA. From 2019 to 2020, he was a Visiting Scholar with the Robert W. Galvin Center for Electricity Innovation, Illinois Institute of Technology, Chicago, IL, USA. He is currently a Professor with the State Key Laboratory of Electrical Insulation and Power Equipment, School of Electrical Engineering, Xi'an Jiaotong University, Xi'an, China. His research interests include electricity markets, power system economics and optimization methods, and power system planning and reliability evaluation. He is an Editor of *IEEE Transactions on Power Systems*, *IEEE Power Electronics Letters*, *IET Generation, Transmission and Distribution*, and *CSEE Journal of Power and Energy Systems*.