2024 3rd Asian Conference on Frontiers of Power and Energy (ACFPE) | 979-8-3503-6767-6/24/$31.00 ©2024 IEEE | DOI: 10.1109/ACFPE63443.2024.10801026

Soft Actor Critic-assisted Receding Horizon Generation Redispatch for Power System Resilience Enhancement against Extreme Weather

Xiangyang Guo
Northwest Branch of State Grid Corporation of China
Xi'an, China

Ziyue Dang
Northwest Branch of State Grid Corporation of China
Xi'an, China

Wujing Li
Northwest Branch of State Grid Corporation of China
Xi'an, China

Biao Su
Northwest Branch of State Grid Corporation of China
Xi'an, China

Xin Liu
Northwest Branch of State Grid Corporation of China
Xi'an, China

Jiahao Wang
Northwest Branch of State Grid Corporation of China
Xi'an, China

Zuibing Xie
Northwest Branch of State Grid Corporation of China
Xi'an, China

Longxiang Duan
College of Electrical Engineering Sichuan University
Chengdu, China
duanlongxiang@stu.scu.edu.cn

***Abstract*—Extreme weather-related power outages occur more frequently in the worldwide due to the climate change. In this paper, we focus on the receding horizon control during real-time operation in order to enhance the power system resilience against the impending extreme weather. The mathematic formulation for receding horizon generation redispatch is firstly presented. Considering this optimization problem as a Markov decision process, the soft actor critic (SAC)-based deep reinforcement learning algorithm is proposed to enable the real-time decision-making for power system resilience enhancement against the extreme weather events. Case study is presented to demonstrate that the proposed SAC-based strategy can reduce the loss of load and adapt to the uncertain scenarios under the impact of the extreme weather events.**

***Keywords*—Extreme weather event, power system resilience, generation redispatch, deep reinforcement learning, receding horizon control.**

I. INTRODUCTION

Due to climate change, extreme weather events, such as hurricanes and ice storms, occur more frequently and also higher intensity. The impacts of extreme weather events on power system is also becoming more and more significant. According to [1], it is estimated that the economic losses due to extreme weather-related power outages is more than $25 billion per year. In January 2008, 451 overhead lines were out due to the damage of tower by the ice storm in Southern China power grid [2]. In August 2017, hurricane Harvey caused about 300 000 customer outages in Texas [3]. Although power systems are usually designed to be robust to *N*-1 or even some *N*-2 contingencies, extreme weathers may damage multiple components in the affected area. As such, power system resiliency against extreme weather events are receiving more and more attention in power engineering society.

In [4], the definition of power system resilience is discussed and a time-series simulation model is proposed for resilience evaluation. In [5], the temporal failure probability of transmission towers and conductors are proposed considering the track of the hurricane and the energy-not-supplied (ENS) index is computed based on Monte Carlo simulation to evaluate the impact of hurricane.

To prevent power system from large-scale outage, both long-term and short-term measures are proposed to enhance the power system resiliency. For long-term measures, facilities reinforcement and transmission expansion planning are studied. In [6], critical lines are selected to be placed underground based on the proposed stochastic robust optimization model for strengthening transmission systems against extreme weather events characterized by high wind speed, such as hurricanes and tornadoes. In [7], line hardening is proposed considering the worst-case disruption scenario and the worst-case wind generation.

On the other hand, short-term measures, including preventive dispatching, emergency dispatching and also restorative dispatching, are proposed to enhance the resilience against an impending extreme weather event. In [8], a wind loading-related transmission tower fragility model is developed to generate the probabilistic operating scenarios considering extreme wind conditions and a stochastic security constrained unit commitment model is proposed to minimize the cost for generation and load shedding. In [9], day-ahead unit commitment under extreme weather events is formulated as a two-stage robust optimization problem and then is solved by using the Column & Constraint Generation algorithm so as to enhance the power system against the worst damage scenario. In [10], a resilient unit commitment model is proposed to coordinating three-stage controlling measures including

The work is supported by Science and Technology Project of Northwest Branch of State Grid (Grant No.: SGNW0000DKJS2310090(B1)).

preventive unit commitment, emergency load shedding, and transmission line restoration.

In this paper, we focus on the receding horizon control during real-time operation in order to enhance the power system resilience[11] against the impending extreme weather[12]. This receding horizon generation redispatch problem can be considered as a Markov decision process. However, the receding horizon generation redispatch problem is related with dynamic programming. It is challenging to compute the solution at each dispatching period. The success of deep reinforcement learning in robotic control and many other research field has provided a prospective alternative. In this paper, to enable real-time decision making, the soft actor critic-based reinforcement learning algorithm is proposed to generate the control action of preventive generation redispatch.

The rest of this paper is organized as follows. In Section II, the soft actor critic algorithm is briefly introduced. In Section III, the proposed soft actor critic-based strategy for receding horizon generation re-dispatching is proposed. In Section IV, case study is presented. In Section V, conclusion is made.

II. Preliminaries of Soft Actor Critic

Soft actor critic (SAC) is an off-policy algorithm that is trained to maximize the trade-off between the expected return and the entropy. It consists of three neural networks, namely a state value function V_ψ parameterized by ψ, a soft Q-function Q_θ parameterized by θ, and a policy function π_ϕ parameterized by ϕ [13].

The soft value function V_ψ is trained to minimize the squared residual error $J_V(\psi)$ as is in (1):

$$J_V(\psi) = \mathbb{E}_{s_t \sim \mathcal{D}}\left[\frac{1}{2}\left(V_\psi(s_t) - \mathbb{E}_{a_t \sim \pi_\phi}\left[Q_\theta(s_t, a_t) - \log \pi_\phi(a_t|s_t)\right]\right)^2\right] \quad (1)$$

where s_t and a_t are the state and the action at stage t. $\mathbb{E}$ denotes the computation of expectation. $\mathcal{D}$ denotes the experience replay buffer.

The Q network Q_θ is trained by minimizing the following loss:

$$J_Q(\theta) = \mathbb{E}_{(s_t,a_t)\sim\mathcal{D}}\left[\frac{1}{2}\left(Q_\theta(s_t,a_t) - \hat{Q}(s_t,a_t)\right)^2\right] \quad (2)$$

where

$$\hat{Q}(s_t,a_t) = r(s_t,a_t) + \gamma\mathbb{E}_{s_{t+1}\sim p}\left[V_{\bar{\psi}}(s_{t+1})\right] \quad (3)$$

$V_{\bar{\psi}}$ is the target value function that shares the same structure with the soft value function V_ψ. But the model parameters of $V_{\bar{\psi}}$ is updated by (4):

$$\bar{\psi} \leftarrow \tau\psi + (1-\tau)\bar{\psi} \quad (4)$$

where τ is the learning rate coefficient.

The policy network π_ϕ is trained by minimizing the Kullback-Leibler divergence between the distribution of the policy function and the distribution of the exponentiation of the Q function normalized by another function Z as in (5):

$$J_\pi(\phi) = \mathbb{E}_{s_t\sim\mathcal{D}}\left[\mathrm{D}_{\mathrm{KL}}\left(\pi_\phi(\cdot\,|s_t)||\frac{\exp(Q_\theta(s_t,\cdot))}{Z_\theta(s_t)}\right)\right] \quad (5)$$

where $\mathrm{D}_{\mathrm{KL}}(x||y)$ denotes the Kullback-Leibler divergence between two distribution *x* and *y*.

III. The Proposed SAC-based Strategy for Generation Re-dispatching

In this section, the mathematical formulation for receding horizon control against extreme weather event is proposed. After that, the SAC-based strategy for receding horizon decision-making is proposed.

A. Optimization Objective

In the proposed scheme, the objective is to minimize the combined cost of generation redispatch and load shedding as in (6):

$$\min \sum_{t=t_0+1}^{T}\left[\sum_{g\in\Omega_G}\left(c_g^{G,UP}\Delta P_{g,t}^{G,UP} + c_g^{G,DW}\Delta P_{g,t}^{G,DW}\right) + \sum_{l\in\Omega_L} c_l^L \Delta P_{l,t}^L\right] \quad (6)$$

where Ω_G denotes the set of power plants. $\Delta P_{g,t}^{G,UP}$ is the upward adjustment of real power generation of the *g*th power plant at time *t* while $c_g^{G,UP}$ is the corresponding cost coefficient. $\Delta P_{g,t}^{G,DW}$ is the downward adjustment of real power generation of the *g*th power plant at time *t* and $c_g^{G,DW}$ is also the corresponding cost coefficient. Ω_L denotes the set of loads. $\Delta P_{l,t}^L$ is the amount of load shedding at time *t*. ΔP_g^G is the real power generation adjustment of the *g*th power plant and c_g^G denotes the corresponding cost coefficient. t_0 denotes the current stage and *T* denotes the maximum stage that the extreme weather event is ended.

B. Constraints

During generation redispatch at time *t*, constraints including power balance, power flow, the limits of active power generation, the ramping rate at each time interval, the power flow security of each transmission line should all be satisfied. In the following, the constraints are formulated.

1) Power Balance Constraints.

At each time instance *t*, the power balance at each bus *n* should be satisfied:

$$\sum_{g\in\Omega_n^G} P_{g,t}^G - \left(P_{n,t}^L - \Delta P_{n,t}^L\right) + \sum_{m\in\Omega_n^{Nbr}} P_{n,m,t}^{Br} = 0$$

$$\forall n \in \Omega_B, t \in [t_0, t_0+1, \cdots, T] \quad (7)$$

where Ω_n^G is the set of generators that is connected to bus *n* and $P_{g,t}^G$ is the active power output of generator *g* at time *t*. $P_{n,t}^L$ is the load demand at bus *n* at state $s_{i,t}$ while $\Delta P_{n,t}^L$ is the corresponding load curtailment at bus *n*. Ω_n^{Nbr} is the set of branches that is connected to bus *n* and $P_{n,m,t}^{Br}$ denotes the active power that is transmitted from bus *n* to the neighboring bus *m*.

2) Operating Constraints of Generators.

For each generator, the upper and lower limits of its real power output at each stage should be satisfied as in (8-9):

$$P_{g,t}^{G} = P_{g,t}^{G,0} + \Delta P_{g,t}^{G,UP} - \Delta P_{g,t}^{G,DW} \tag{8}$$

$$\underline{P_g^G} \cdot o_t \leq P_{i,t,g}^{G} \leq \overline{P_g^G} \cdot o_t, g \in \Omega^G \tag{9}$$

where $\underline{P_g^G}$ and $\overline{P_g^G}$ are the minimum and maximum real power output of generator *g*.

Apart from the real power output constraint, the ramping rate constraint between two successive stage should be satisfied as in (10-11):

$$P_{g,t+1}^{G} - P_{g,t}^{G} \leq R_g^{up} \tag{10}$$

$$P_{g,t}^{G} - P_{g,t+1}^{G} \leq R_g^{down} \tag{11}$$

where R_g^{up} and R_g^{down} are the maximum ramping up and down limits for generator *g*.

3) Branch Power Flow and its Security Constraints.

The limit for power flows through the online branches should be satisfied:

$$B_{n,m}\left(\theta_{n,t} - \theta_{m,t}\right) - P_{n,m,t}^{Br} + \left(1 - u_{n,m,t}\right)M \geq 0$$

$$\forall \langle n,m \rangle \in \Omega^{Br} \tag{12}$$

$$B_{n,m}\left(\theta_{n,t} - \theta_{m,t}\right) - P_{n,m,t}^{Br} + \left(1 - u_{n,m,t}\right)M \leq 0$$

$$\forall \langle n,m \rangle \in \Omega^{Br} \tag{13}$$

$$\underline{P_{n,m}^{Br}} \cdot u_{n,m,t} \leq P_{n,m,t}^{Br} \leq \overline{P_{n,m}^{Br}} \cdot u_{n,m,t}$$

$$\forall \langle n,m \rangle \in \Omega^{Br} \tag{14}$$

where θ represents the phase angle of bus voltage. $B_{n,m}$ is susceptance of branch $\langle n,m \rangle$ in DC power flow computation. $u_{n,m,t}$ is the operating status of branch $\langle n,m \rangle$. $u_{n,m,t} = 1$ indicates branch $\langle n,m \rangle$ is operating while $u_{n,m,t} = 0$ indicates branch $\langle n,m \rangle$ is out-of-service due to the damage by windstorm. *M* is a large constant for the big-M method.

4) Limit for Voltage Phase Angle in DC Power Flow.

$$\underline{\theta_n} \leq \theta_n \leq \overline{\theta_n}, \forall n \in \Omega_B \tag{15}$$

5) Limit for Decision Variables.

$$\Delta P_{g,t}^{G,UP} \geq 0 \tag{16}$$

$$\Delta P_{g,t}^{G,DW} \geq 0 \tag{17}$$

$$0 \leq \Delta P_{l,t}^{L} \leq P_{l,t}^{L} \tag{18}$$

The mathematical problem defined by (6-18) can be sorted as a dynamic programming problem. During real-time operation, this optimization problem is solved repeatedly in the manner of receding horizon control. In other words, at each dispatching stage, the optimization problem is solved and the preventive control actions from the current stage to the ending stage of the extreme weather can be obtained. After that, the actions for the current stage are implemented while the other actions for the future stage will not be necessary implemented. This above-mentioned procedure is repeated as so to make the preventive control actions at each stage to be beneficial to both the current stage and future stage under uncertain operating scenarios that are characterized by the damages of transmission lines and the subsequent changes of network topologies.

C. The Markov Decision Process for Preventive Control against Extreme Weather

The receding horizon control against extreme weather event can be considered as a Markov decision process (MDP). The environment, the state, the action and the reward should be defined as following.

1) The Environment E.

The power system can be considered as the interactive environment for the MDP.

2) The State S.

The state reflects the operating condition of the power system at each dispatching instant during the development of an extreme weather event. The input variables for power flow computation, including the load demand, the generation and the network topology, are used to represent the state. In this case, the state $\mathcal{S}$ can be formulated as in (19):

$$\mathcal{S}(t) = \left[u_{1,t}^{BR}, \cdots, u_{N_{BR},t}^{BR}, P_{1,t}^{G}, \cdots, P_{N_G,t}^{G}, P_{1,t}^{L}, \cdots, P_{N_L,t}^{L}\right] \tag{19}$$

3) The action A.

The action is related with the decision variables in (16-18). So the action consists of the regulation of the active power outputs of generators and the load shedding:

$$\mathcal{A}(t) = \left[\Delta P_{1,t}^{G,UP}, \cdots, \Delta P_{N_G,t}^{G,UP}, \Delta P_{1,t}^{G,DW}, \cdots, \Delta P_{N_G,t}^{G,DW}, \Delta P_{1,t}^{L}, \cdots, \Delta P_{N_L,t}^{L}\right] \tag{20}$$

4) The reward R.

In deep reinforcement learning, the reward is used to guide the neural network-based agent to optimize its policy. In this paper, the goal is to minimize the cost of generation redispatch and load shedding. So it is intuitive to define the reward as the reverse of the objective function in (5). However, as the problem is constrained optimization, the violation of any constraints should indicate that the action is infeasible. In this regard, penalties should be introduced into the reward function so as to prevent the infeasible actions. The reward function is defined as in (21):

$$\mathcal{R}(t) = \mathcal{R}_{obj} + \mathcal{R}_{ctr} \tag{21}$$

where

$$\mathcal{R}_{obj} = exp\big(-\textstyle\sum_{g\in\Omega_G}\big(c_g^{G,UP}\Delta P_{g,t}^{G,UP} + c_g^{G,DW}\Delta P_{g,t}^{G,DW}\big) + \textstyle\sum_{l\in\Omega_L} c_l^L \Delta P_{l,t}^L\big) \tag{22}$$

$$\mathcal{R}_{ctr} = \textstyle\sum_{i\in\Omega_{ctr}} \mathcal{P}_i^{ctr} \tag{23}$$

$$\mathcal{P}_i^{ctr} = \begin{cases} -1, if\ inequality\ contraint\ is\ violated \\ 0, if\ inequality\ contraint\ is\ satisfied \end{cases} \tag{24}$$

Here, Ω_{ctr} denotes the set of inequality constraints in (9-15). $\mathcal{P}_i^{ctr}$ is the penalty that is related with the ith inequality constraint.

D. The Proposed SAC-based Reinforcement Learning Strategy

After defining the MDP, the SAC algorithm is used to train the agent for decision-making for power system resilience enhancement. The proposed strategy consists of two stages, namely the offline agent training stage and the online decision-making stage. The procedures for these two stages are explained as following.

In offline agent training stage, the neural network-based agents learn to fine tune their own parameters by interacting with the environment. At each training epoch, the procedure is as following:

1) Start the interactive procedure and set $t = t_0$.

2) Form the state vector $\mathcal{S}(t)$ by gathering the status of transmission lines, the scheduled real power generations and the forecasted load demands.

3) Obtain the action vector $\mathcal{A}(t)$ by the policy network π_ϕ.

4) Implement the dispatching action $\mathcal{A}(t)$. By performing power flow computation, obtain the new state vector $\mathcal{S}'$ and compute the reward $\mathcal{R}(t)$.

5) Store the new instance $\{\mathcal{S}(t), \mathcal{A}(t), \mathcal{R}(t), \mathcal{S}'\}$ into the experience replay buffer $\mathcal{D}$.

6) Let $t = t + 1$. If t is higher than the maximum dispatching instant T, go to 7); otherwise, go back to 2).

7) Start the model training procedure and get a minibatch samples from the experience replay buffer $\mathcal{D}$.

8) Compute the loss by (1) and update the parameters ψ of the soft value function V_ψ by back propagation-based optimizer.

9) Compute the loss by (2) and update the parameters θ of the Q function Q_θ by optimizer.

10) Compute the loss by (5) and update the parameters ϕ of the policy function π_ϕ by optimizer.

11) Update the parameters $\bar{\psi}$ of the target value function by (4).

12) Determine whether the training epoch reaches the maximum iteration. If so, end the procedure and output the policy function π_ϕ; otherwise, go back to 1).

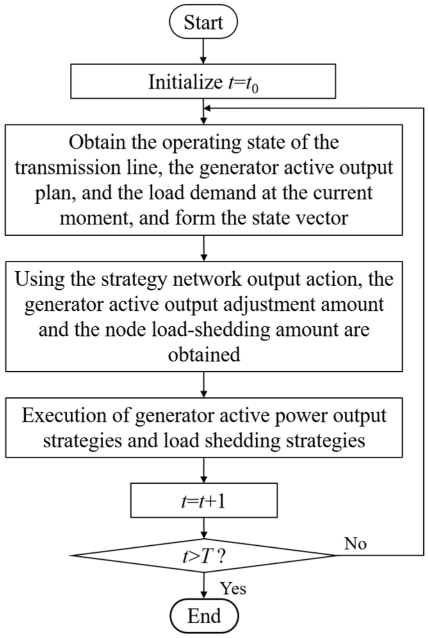
Start
Initialize t=t_0
Obtain the operating state of the transmission line, the generator active output plan, and the load demand at the current moment, and form the state vector
Using the strategy network output action, the generator active output adjustment amount and the node load-shedding amount are obtained
Execution of generator active power output strategies and load shedding strategies
t=t+1
t>T ?
No
Yes
End

Fig. 1.   Flowchart of online generation redispatch

In online decision-making stage, at each dispatching instant, together with the operating status of all the transmission lines, the scheduled real power generations and the forecasted load demands are collected to form the state $\mathcal{S}(t)$. Then the action $\mathcal{A}(t)$ is obtained by feeding the state $\mathcal{S}(t)$ into the policy network π_ϕ. The dispatching action for instant t, including the adjusted real power generations and the load shedding, are determined by $\mathcal{A}(t)$. Implement the dispatching action and wait for the next dispatching instant. The above procedure is repeated until the extreme weather event is ended.

IV. Case Study

The IEEE RTS 24-node system is used to verify the effectiveness of the proposed strategy. The topology of the IEEE RTS 24-node system and the ice storm intrusion path are shown in Fig. 2. It can be seen that six lines of the grid are located within the affected zone of the ice storm track. According to the coupling model of the ice storm intensity and the ice and wind loads imposed on the lines, it is obvious that the probability of failure of these lines increases significantly. Their failure probabilities are shown in Table 1. Other detailed parameters of the IEEE RTS 24-node system are in Matpower 4.1.

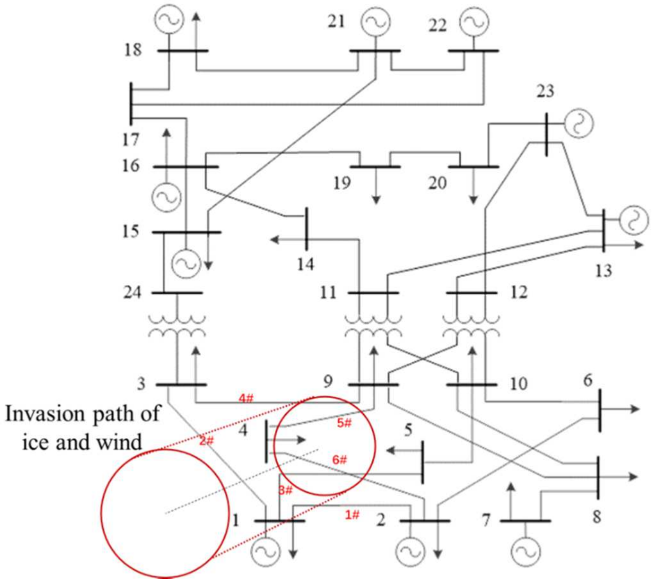

Fig. 2. Ice wind intrusion path and IEEE RTS 24-node system topology

TABLE I. PROBABILITY OF LINE FAILURE

Line No.	Line	Probability of line failure
1#	2-1	0.25
2#	1-3	0.22
3#	1-5	0.2
4#	3-9	0.18
5#	4-9	0.15
6#	2-4	0.12

As the ice storm progresses, transmission lines may fail consecutively. If there is a line in the grid that is close to the other, that line may also fail at the next time section, creating a different fault state. The fault states are related to the geographic location of the line and the path of ice storm intrusion. The fault states and the transfer probabilities between states are shown in Fig. 3.

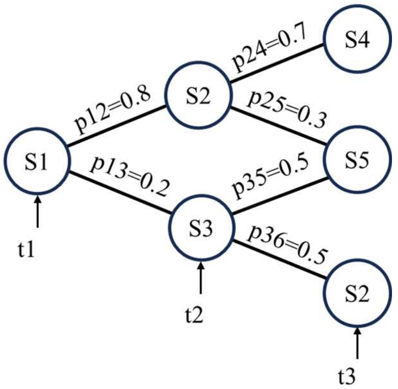

Fig. 3. Fault state transfer probability

According to equation (25), the probability of each state can be calculated as shown in Table 2.

$$P_i^{(PDF)} = \frac{1}{N_T}\sum_{r\in\Omega_i^{Path}} \left(p_{r,i,t}\right) \qquad (25)$$

TABLE II. PROBABILITY OF EACH FAULT STATE

State	Probability
S1	0.33
S2	0.27
S3	0.07
S4	0.19
S5	0.11
S6	0.03

With the time lapse and state transfer, different fault scenarios are formed through different state transfer paths, and the lines with faults under each scenario in each time period are shown in Table 3, with an interval of 1 hour for each time section. Due to the short time scale, the repair of the faulted line is not considered. To simplify the analysis, only four hours of dispatching are considered in this paper, in which no fault occurs at t_0.

TABLE III. FAULT LINES IN EACH SCENARIO

Scenario No.	Fault lines			
	Time t_0	***Time t_1***	***Time t_2***	***Time t_3***
1	No	1#	1# 2#	1# 2# 4#
2	No	1#	1# 2#	1# 2# 4# 5#
3	No	1#	1# 3#	1# 3# 6#
4	No	1#	1# 3#	1# 3# 5# 6#

Based on the strategy proposed in this paper, after the possible fault scenarios and their occurrence probabilities are obtained from the computational model. The unit output is adjusted as a precautionary control decision before the ice storm affects the grid, in order to minimize the loss of load due to line faults, and to reduce the risk of the grid caused by extreme weather. The adjusted amount of each unit output at different time sections for the four scenarios is shown in Fig. 4.

As can be seen in Figure 4, before a possible line break fault occurs in the grid, some generating units pre-adjust their output to reduce the risk of load loss in the subsequent hours while meeting the constraints of the grid line capacity, the upper limit of unit output, and the node phase angle. Among them, although G23 and G24 are farther away from the risk area, their output adjustments are larger and play an important role in load preservation by supporting the affected nodes through the contact line.

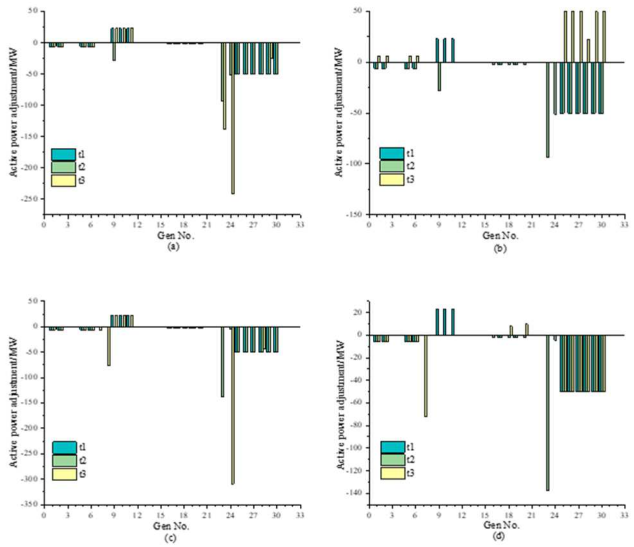

Fig. 4. Adjustment of unit output in each scenario

As can be seen in Fig. 5, the preventive control strategy proposed in this paper for adjusting unit output can effectively reduce the amount of lost load in fault scenarios to minimize the impact of extreme ice storms on the grid, compared to no preventive strategy.

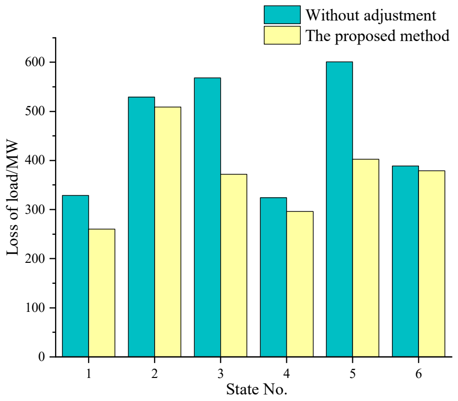

Fig. 5. Comparison of loss of load

V. Conclusion

In this paper, we focus on the receding horizon control during real-time operation in order to enhance the power system resilience against the impending extreme weather. The mathematic formulation for receding horizon generation redispatch is firstly presented. Considering this optimization problem as a Markov decision process, the soft actor critic (SAC)-based deep reinforcement learning algorithm is proposed to enable the real-time decision-making for power system resilience enhancement against the extreme weather events. Case study is presented to demonstrate that the proposed SAC-based strategy can reduce the loss of load and adapt to the uncertain scenarios under the impact of the extreme weather events.

Acknowledgment

The work is supported by Science and Technology Project of Northwest Branch of State Grid (Grant No.: SGNW0000DKJS2310090(B1)).

References

[1] Campbell, R. J. . "Weather-Related Power Outages and Electric System Resiliency [August 28, 2012]." Congressional Research Service Reports (2012).

[2] Ping-yuan Liu, Hong-ming He and Chun-ping Pan, "Investigation of 2008 frozen disaster and research on de-icing in Guangdong power grid," 2008 China International Conference on Electricity Distribution, Guangzhou, China, 2008, pp. 1-5.

[3] US Dept. Energy, Washington, DC, USA, “Infrastructure security and energy restoration, Hurricane Harvey event report (update 4), 2017.

[4] M. Panteli and P. Mancarella, "Modeling and Evaluating the Resilience of Critical Electrical Power Infrastructure to Extreme Weather Events," IEEE Systems Journal, vol. 11, no. 3, pp. 1733-1742, Sept. 2017.

[5] H. Zhang, L. Cheng, S. Yao, T. Zhao and P. Wang, "Spatial–Temporal Reliability and Damage Assessment of Transmission Networks Under Hurricanes," IEEE Transactions on Smart Grid, vol. 11, no. 2, pp. 1044-1054, March 2020.

[6] D. N. Trakas and N. D. Hatziargyriou, "Strengthening Transmission System Resilience Against Extreme Weather Events by Undergrounding Selected Lines," IEEE Transactions on Power Systems, vol. 37, no. 4, pp. 2808-2820, July 2022.

[7] A. Bagheri, C. Zhao, F. Qiu and J. Wang, "Resilient Transmission Hardening Planning in a High Renewable Penetration Era," IEEE Transactions on Power Systems, vol. 34, no. 2, pp. 873-882, March 2019.

[8] Y. Sang, J. Xue, M. Sahraei-Ardakani and G. Ou, "An Integrated Preventive Operation Framework for Power Systems During Hurricanes," IEEE Systems Journal, vol. 14, no. 3, pp. 3245-3255, Sept. 2020.

[9] D. N. Trakas and N. D. Hatziargyriou, "Resilience Constrained Day-Ahead Unit Commitment Under Extreme Weather Events," in IEEE Transactions on Power Systems, vol. 35, no. 2, pp. 1242-1253, March 2020.

[10] T. Ding, M. Qu, Z. 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.

[11] H. Guo, L. Chen, Q. Zhang, H. Huang, Q. Ma and J. Wang, "Research and Response to Extreme Scenarios in New Power System: A Review from Perspective of Electricity and Power Balance," Power System Technology, pp. 1-27, September 2024.

[12] "Improving wind power utilisation under stormy weather condition by risk-limiting unit commitment," IET Renewable Power Generation, vol 12, no. 15, pp. 1778-1785, November, 2018

[13] Haarnoja, Tuomas, et al. "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor." International conference on machine learning. PMLR, 2018.