A Data-driven Shunt Dispatch Approach to Enhance Power System Resilience against Windstorms

MD Kamruzzaman[1], Michael Abdelmalak[1], Salem Elsaiah[2], and Mohammed Benidris[1]
[1]Department of Electrical and Biomedical Engineering, University of Nevada-Reno, Reno, NV 89557, USA
[2]School of Engineering, State University of New York, Maritime College, Throggs Neck, NY 10465
Emails: {mkamruzzaman, mabdelmalak}@nevada.unr.edu, selsaiah@sunymaritime.edu, mbenidris@unr.edu

***Abstract*—This paper proposes a machine learning-based approach to dispatch shunt reactive power compensators for enhancing operational resilience of power systems against windstorms. Existing resilience enhancement approaches do not dispatch shunts to maintain voltage magnitudes within the standard limits during extreme weather events, and are computationally expensive. In this work, a data-driven approach is proposed to both exploit the capability of shunts in enhancing power system operational resilience against windstorms and to overcome the computational burden of analytical optimization methods. In the proposed approach, a multi-agent framework is developed using the fundamental continuous soft actor critic algorithm to dispatch shunts during windstorms. The developed multi-agent framework is trained using hypothetical multiple line outages, which are generated based on line fragility curves against windstorms. The trained network dispatches shunt compensators whenever there are voltage violations during windstorms (i.e., dependent multiple line outages in the path of a windstorm), thereby enhancing the operational resilience of power systems against windstorms. The proposed method is demonstrated on the IEEE 30-Bus system. The results show that voltage magnitudes can be maintained within the standard limits for most of simulated scenarios.**

***Index Terms*—Deep reinforcement learning, power system resilience, shunt reactive power, soft-actor-critic algorithm.**

I. Introduction

The frequency and intensity of extreme weather events have imposed significant impacts on the performance of power grids resulting in prolonged power outages and economic losses [1], [2]. Various resilience enhancement strategies have been proposed in the literature to reduce, mitigate, and prevent the impacts of weather events on performance of power systems. These strategies include mobile energy storage devices, topology switching, and load shedding. Nonetheless, probabilistic strategies that encounter the stochastic behavior of extreme events, in particular windstorms, have not gained much interest. In addition, challenges that face providing a fast acting enhancement algorithm, which provides an immediate action to maintain voltage magnitudes within predefined limits, need to be addressed. Furthermore, neglecting impacts of future potential failures, diverse failure scenarios, and infeasible steady-state operation may yield less realistic models and results [1]. Consequently, it has become important to develop a resilience enhancement strategy that provides a proactive voltage control prior to or during extreme weather events.

Numerous power system resilience enhancement methods have been proposed from different standpoints [3]–[14]. In [3], a mixed-integer linear programming (MILP)-based method has been proposed to enhance power system resilience through re-dispatching generators, re-configuring network topology, and shedding loads. An algorithm for enhancing the resilience of a multi-microgrid system via dispatching of unused capacitor banks has been proposed in [4]. Additionally, an MILP-based generation re-dispatch strategy has been proposed in [5] to enhance power system resilience during ice storms. Moreover, several preventive action-based strategies such as a multi-sensor prediction-based wide-area monitoring and control [11], a linear-programming-based optimal siting and sizing of energy storage devices [12], a Monte-Carlo simulation (MC)-based proactive unit commitment framework [13], and an MC-based crew preposition and network reconfiguration technique [14] have been proposed to enhance power system resilience.

It is important to highlight here that the methods proposed in [3]–[14] are effective means to enhance the resilience of power systems. However, these methods do not fully explore the potential of using shunt compensators to enhance power system resilience. More significantly, these methods are not computationally flexible to enhance operational resilience of power systems due to their dependency on accurate system information. Therefore, a resilience enhancement method that does not require accurate system knowledge and flexible to control shunt compensators output power needs to be developed to enhance operational resilience of power grids.

This paper proposes a machine learning-based approach to enhance power system resilience against windstorms. The proposed method is developed based on dispatching of reactive power compensators, and thereby preserving bus voltages within the acceptable limits in case of single or multiple line outages. The fundamental continuous soft actor critic (CSAC) algorithm is used to develop a multi-agent framework to control the reactive power output of shunt compensators. In the proposed method, power systems are divided into regions, where each region represents an agent. The algorithm is trained using historical data and fragility curves of transmission lines against windstorms. The trained algorithm is then used to provide corrective control actions when a power system is impacted by a windstorm. The proposed algorithm is tested on several systems including the IEEE 30-bus system.

The rest of the paper is organized as follows. Section II describes the proposed approach. Section III illustrates the proposed algorithm. Section IV provides case studies. Section V provides concluding remarks.

II. The Proposed Multi-agent Framework

A. Policies for Actor-networks of the Proposed Framework

Each agent in the proposed multi-agent framework has one actor network to provide actions, which is developed using a squashed Gaussian distribution function. The policy of the actor network to provide actions is expressed as follows.

$$\alpha_t^{ci} \sim \pi_{\xi^{ci}}(\alpha_t^{ci}|O_t^i), \tag{1}$$

where i represents the i^{th} agent of the multi-agent framework; O_t^i is the observation vector of the i^{th} agent at time t; α_t^{ci} is the provided action by the actor-network of the i^{th} agent; ξ^{ci} is the parameter for policy of the i^{th} agent; and $\pi_{\xi^{ci}}(\alpha_t^{ci}|o_t^i)$ is an unbounded Gaussian policy of the i^{th} agent. A squashing function need to be applied on $\pi_{\xi^{ci}}(\alpha_t^{ci}|o_t^i)$ to bound actions of the i^{th} agent to a finite value.

B. Policy Training Algorithm for Actors

In the fundamental CSAC algorithm, the policy is updated in each iteration to maximize the expected return and entropy (randomness measure of the policy). Following the same convention, policies of the proposed algorithm are also updated in each iteration. A value function, $V_{\psi^i}^{ci}(O_t^i)$, which is used to measure the soft value for policy of the i^{th} agent can be expressed as follows.

$$V_{\psi^i}^{ci}(O_t^i) = \mathop{\mathbb{E}}_{\alpha_t^{ci}\sim\pi_{\xi^{ci}}}\left[Q_\theta(s_t,\alpha_t^{ci},\alpha_t^{-ci}) - \alpha^{ci}\log\left(\pi_{\xi^{ci}}(\alpha_t^{ci}|O_t^i)\right)\right] \tag{2}$$

where ψ^i represents parameter of the value function network for the i^{th} agent; θ represents parameter for the Q value function; $Q_\theta(s_t,\alpha_t^{ci},\alpha_t^{-ci})$ is a critic or centralized policy evaluation function for all the actors; α_t^{-ci} is the action provided by actors of agents except agent i; α^{ci} represents a parameter to determine the relative importance between reward and entropy of the i^{th} agent; and s_t is a set for system states.

The expression provided in (3) is used to minimize the residual squared error of a soft Bellman function to train value functions of the actors.

$$J_v^{ci}(\psi^i) = \mathbb{E}_{s_t^{ci}\sim\mathcal{D}}\left[\frac{1}{2}V_{\psi^i}^{ci}(O_t^i) - \left[Q_\theta(s_t,\alpha_t^{ci},\alpha_t^{-ci}) - \alpha^{ci}\log\left(\pi_{\xi^{ci}}(\alpha_t^{ci}|O_t^i)\right)\right]^2\right] \tag{3}$$

where $\mathcal{D}$ is a replay buffer to store experiences of the actors.

The gradient of (3) using an unbiased estimator is determined as follows to sample actions from the current policy.

$$\hat{\nabla}_{\psi^i}J_v^{ci}(\psi^i) = \nabla_{\psi^i}V_{\psi^i}^{ci}(O_t^i)\left(V_{\psi^i}^{ci}(O_t^i) - Q_\theta(s_t,\alpha_t^{ci},\alpha_t^{-ci}) + \alpha^{ci}\log\left(\pi_{\xi^{ci}}(\alpha_t^{ci}|O_t^i)\right)\right) \tag{4}$$

In this work, we have modified the expression for training the soft-Q parameters of the basic actor given in [15], which can be expressed as follows.

$$J_{Q_\theta}^{ci}(\theta^i) = \mathbb{E}_{(s_t^{ci},\alpha_t^{ci})\sim\mathcal{D}}\left[\frac{1}{2}\left(Q_\theta(s_t,\alpha_t^{ci},\alpha_t^{-ci}) - \hat{Q}(s_t,\alpha_t^{ci},\alpha_t^{-ci})\right)^2\right] \tag{5}$$

where

$$\hat{Q}(s_t,\alpha_t^{ci},\alpha_t^{-ci}) = r(s_t,\alpha_t^{ci},\alpha_t^{-ci}) + \beta\mathbb{E}_{s_{t+1}\sim p}[V_{\bar{\psi}^i}^{ci}(o_{t+1}^i)] \tag{6}$$

and $\beta\in[0,1]$ is a discount factor and $\bar{\psi}^i$ is an average of the weights for the value network of i^{th} agent.

The value of Q-function (5) is optimized as follows.

$$\hat{\nabla}_{\theta^i}J_{Q_\theta}^{ci}(\theta^i) = \nabla_{\theta^i}(Q_\theta(s_t,\alpha_t^{ci},\alpha_t^{-ci})\left(Q_\theta(s_t,\alpha_t^{ci},\alpha_t^{-ci}) - r(s_t,\alpha_t^{ci},\alpha_t^{-ci}) - \beta V_{\bar{\psi}^i}^{ci}(o_{t+1}^i)\right)^2 \tag{7}$$

The policy needs to be updated in each iteration to maximize the rewards for improving the policy. The authors of [15] directed the policy update toward exponential of new soft Q-function as they intended to track the policy update. Also, the potential policies are restricted to a parameterized distributions (i.e., Gaussian) family. Following the same convention, we updated the expression for policy update of basic CSAC algorithm for the proposed algorithm as follows.

$$\pi_{\xi^{ci}}^{new} = \arg\min D_{KL}\left(\pi_{\xi^{ci}}(.|O_t^i)\middle|\middle|\frac{Q_\theta(s_t,.)}{Z_\theta(s_t)}\right) \tag{8}$$

where $Z_\theta(s_t)$ is an intractable partition function, which does not contribute to the gradient with respect to the new policy.

The policy $\pi_{\xi^{ci}}(.|O_t^i)$ is parameterized for action setting using the policy network of agent, i, with parameter ξ^{ci}. Finally, the expected KL-divergence of (8) is multiplied by α^{ci}, and then, minimized ignoring $Z_\theta(s_t)$ to train the policy parameters of agent, i, as follows.

$$J_{\pi_{\xi^{ci}}}^{ci}(\xi^{ci}) = \mathbb{E}_{s_t^{ci}\sim\mathcal{D}}\left[\mathbb{E}_{\alpha_t^{ci}\sim\pi_{\xi^{ci}}}\left[\alpha^{ci}\log\left(\pi_{\xi^{ci}}(\alpha_t^{ci}|O_t^i)\right) - Q_\theta(s_t,\alpha_t^{ci},\alpha_t^{-ci})\right]\right] \tag{9}$$

Although several options are available to minimize the objective function $J_{\pi_{\xi^{ci}}}^{ci}(\xi^{ci})$, the authors of [16] have applied the reparameterization trick to achieve target density which is Q-function. The modified expression to reparameterize the policy of agent, i, is as follows.

$$\alpha_t^{ci} = f_{\xi^{ci}}(\epsilon_t^{ci};o_t^i) \tag{10}$$

where ϵ_t^{ci} is a noise vector that is using a spherical Gaussian distribution.

Thus, the new policy objective for agent, i, is as follows.

$$J_{\pi_{\xi^{ci}}}^{ci}(\xi^{ci}) = \mathbb{E}_{s_t^{ci}\sim\mathcal{D},\epsilon_t^{ci}\sim\mathcal{N}}\left[\alpha^{ci}\log\left(\pi_{\xi^{ci}}(f_{\xi^{ci}}(\epsilon_t^{ci};o_t^i)|O_t^i)\right) - Q_\theta(s_t,f_{\xi^{ci}}(\epsilon_t^{ci};o_t^i),f_{\phi^{-ci}}(\epsilon_t^{-ci};s_t^{-i}))\right] \tag{11}$$

where $f_\phi^{-ci}(\epsilon_t^{-ci};s_t^{-i})$ is the parameterized policies of other actors.

In [17], the authors provided detailed formulation of an alternative approach to obtain the temperature parameter learning objective function, which is not strictly relevant to this work. However, we modify their temperature objective function for the actors of each agent of the proposed framework as follows.

$$J^{ci}(\alpha^{ci}) = \mathbb{E}_{\alpha_t^{ci}}\sim\pi_{\xi^{ci}}\left[-\alpha^{ci}\left(\log\left(\pi_{\xi^{ci}}(\alpha_t^{ci}|O_t^i) + \bar{H}\right)\right]\right. \tag{12}$$

where $\bar{H}$ is an equivalent constant vector of the hyper-parameter to represent target entropy. Equation (12) cannot be minimized directly due to the expectation operator. Therefore, it is minimized using a MC estimator after sampling experiences from replay buffer based on the procedure from [17]. In the proposed multi-agent algorithm, two soft Q-networks for all agents are trained, and then, the minimum value among the outputs of two Q-networks is used in the objective function of (12) to combat state-value overestimation [18].

III. The Proposed Shunt Dispatch Algorithm

A. Modeling of Windstorms

The performance of power system components is directly impacted by the propagation of extreme weather events through the system topology. Most resilience-based studies have used probabilistic fragility curves to model the stochastic behavior of component failures against extreme weather events [19]. Developing a novel algorithm to model windstorms is not germane to the presented work. Consequently, we adopted the provided method in [20] to model windstorms using probabilistic fragility curves. Also, instead of reproducing the provided rigorous mathematical model in [20], we provide only the required expressions to model failures of power system components due to windstorms. During windstorms, component failure rates can be evaluated as follows.

$$\lambda_t^w = [1 + \zeta \cdot (w_t^2/w_{crit}^2 - 1)] \cdot \lambda_{norm} \tag{13}$$

where λ_t^w represents the failure rate at time t during windstorms; w_t is the wind speed in m/s during windstorm; w_{crit} is the critical wind speed in m/s; ζ is the scaling parameter; and λ_{norm} is the failure rate during normal operation.

Weibull distribution is used to model wind speed. The parameters of the Weibull distribution vary according the geographical location based on detailed statistical analysis. On the other hand, critical wind speed defines the threshold upon which a windstorm can be identified. In other words, wind speed below critical value implies normal wind conditions. Also, the scaling parameter varies according the climate and weather fluctuations in specific geographical regions.

B. States, Observations, and Actions

Various parameters can be used to represent system states [21]–[23]; however, for reactive power control studies, voltage magnitudes have been widely used. In this study, voltage states are divided into three zones as shown in Table I.

Each agent is assumed to observe and control voltage profile of the assigned region. Voltage magnitudes are readjusted based on the reactive power output of shunt compensators as well as their locations. The two control variables are continuously updated within their predefined range limits.

C. Calculation of Rewards

A proper reward value, R_k^t, should be defined to assess the effectiveness of the actions. Each agent is encouraged to reduce the deviation of voltage magnitudes during contingencies from a predefined reference value, $V_{ref} = 1.0$ p.u. Rewards could be classified based on voltage operating limits as described in Table I. Generally, the reward value increases as the voltage deviation decreases. If the value of all bus voltages remain in normal or violation zones after dispatching shunts, then the total reward is calculated using (14); otherwise, a relatively large penalty is assigned.

TABLE I. Reward Value based on Voltage Levels

Operation Zone	V_k^t	r_k^t
Normal	$[V_{ref}, V^{ub}]$	$\dfrac{V^{ub} - V_k^t}{V^{ub} - V_{ref}}$
Normal	$[V^{lb}, V_{ref}]$	$\dfrac{V_k^t - V^{lb}}{V_{ref} - V^{lb}}$
Violation	$[V^{ub}, 1.25]$	$\dfrac{V_k^t - V_{ref}}{V_{ref} - 1.25}$
Violation	$[0.8, V^{lb}]$	$\dfrac{V_{ref} - V_k^t}{0.8 - V_{ref}}$
Diverge	$[0.0, 0.8]$	-5
Diverge	$[1.25, \infty]$	-5

$$r^t = \sum_{k=1}^{N^b} R_k^t / N^b \tag{14}$$

D. Training and Execution Algorithms

The power grid is divided into several regions based on the electrical distance between components such that each region is controlled by one agent. The number of regions (agents) varies according to system sizes. The set of bus voltages in each region during a contingency is regulated within the acceptable voltage limits. The set of control actions for each agent, i, can be expressed as follows.

$$\alpha_t^{ci} = \begin{cases} \pi_{\phi^{ci}}(\alpha_t^{ci} | o_t^i), & \text{if } \mid \Lambda_t^i \mid > 0 \\ \alpha_{t-1}^{ci}, & \text{if } \mid \Lambda_t^i \mid = 0 \end{cases} \tag{15}$$

where $\mid \Lambda_t^i \mid$ represents the number of violated bus voltages in the i^{th} region; and α_t^{ci} represents the action specifying the amount and locations of dispatch shunts for the i^{th} agent.

To train all agents, a replay buffer is used as follows.

$$\mathcal{D} \leftarrow \left(s_t, o_t^i, \alpha_t^{ci}, \alpha_t^{-ci}, r^t, s_{t+1}, o_{t+1}^i, \alpha_{t+1}^{ci}, \alpha_{t+1}^{-ci} \right) \tag{16}$$

The training and testing/execution steps for the multi-agent framework are summarized in algorithm 1 and algorithm 2.

IV. Results and Discussion

The proposed approach is applied on a modified IEEE 30-bus system [24]. A windstorm is assumed to pass through the system as shown in Fig. 1. For assessment purpose, we assume that five shunt compensators are located at buses 3, 7, 11, 18, and 27 of the IEEE-30 bus system. Each shunt has a reactive power capacity of 13 MVAr.

To validate the accuracy and effectiveness of the proposed method, the following procedures are implemented sequentially. First, several failure scenarios are created using windstorm modeling approach provided in section III for the defined windstorm in Fig. 1. To capture wide range of failure

Algorithm 1 Training of the Multi-agent Framework

for episode = 1 to M **do**
  Create failure scenario using fragility curve.
  Solve power flow to determine o_i^t and s_t of each agent.
  Count $\mid \Lambda_t^i \mid$ for each agent.
  while voltages violate and step $< N$ **do**
    Evaluate actions, α_t^{ci} for agent i using (15).
    Execute actions α_t^{ci} using power flow solver environment (e.g., Pypower).
    Observe s_{t+1}, r_t, and d to check terminal conditions.
    Store (s_t, o_t^i, α_t^{ci}, α_t^{-ci}, r_t, s_{t+1}, d) in $\mathcal{D}_i$ using (16).
    If s_{t+1} is terminal, reset the environment.
    Update weights of the policies using (11).
    Update the Q-function parameters of local and target networks of each agent using (7).
    Update temperature of actor-networks using (12).
    Update target networks weights of each agent using $\bar{Q}_m \leftarrow \tau Q_m + (1-\tau)\bar{Q}$, where, $m \in \{1,2\}$ and $m \ll 1$.
  end while
end for

Algorithm 2 Testing of the Multi-agent Framework

for episode = 1 to M **do**
  Create failure scenario using fragility curve.
  Solve power flow to determine o_i^t and s_t of each agent.
  Count $\mid \Lambda_t^i \mid$ for each agent.
  while voltages violate and step $< N$ **do**
    Evaluate actions, α_t^{ci} for using (15).
    Execute actions α_t^{ci} using power flow solver.
    Observe s_{t+1}, r_t, and d to validate terminal conditions.
  end while
end for

scenarios, wind speed is assumed to be within 15–51 m/s. For each failure scenario, power flow solution is obtained. Algorithm 1 is used for training the multi-agent framework.

The proposed algorithm is implemented for a fixed number of episodes (failure scenarios). The number of iterations and corresponding rewards for each episode are plotted as shown in Fig. 2a and Fig. 2b, respectively. It is obvious that as the algorithm explores more scenarios, the action time decreases and the reward value increases. The learning rate of agents is enhanced based on previous experiences to avoid bad actions. From Fig. 2a and Fig. 2b, we can see that the ability of agents to resolve the impacts of windstorms on voltage constraints advances very quickly after 15000 episodes and noticeable increase happens in the reward values. For further details, Fig. 3a shows the losses for the critics that fluctuate at the beginning of episodes' period, and finally converge to equilibrium solutions. For accuracy validation, the trained agents are tested using a set of failure scenarios included in testing data. Fig. 3b shows the number of iterations and reward

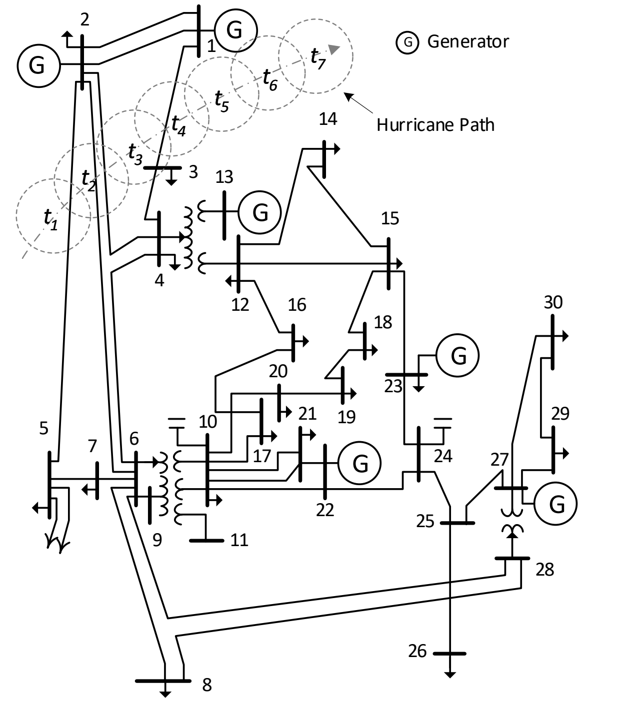
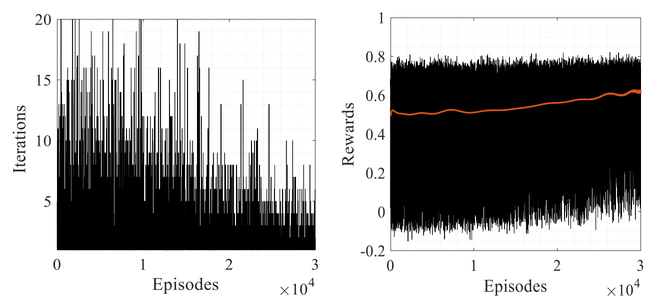


Fig. 1. IEEE 30-bus System.



(a) (b)

Fig. 2. (a) Required number of iterations (b) amount of rewards of training episodes for the IEEE-30 bus system.

values, respectively, for testing data. The trained agents are able to determine proper actions to control shunt compensators within one iteration with maximum reward value for testing scenarios. Thus, the proposed multi-agent framework is trained to provide actions to control shunts reactive power output.

Finally, the trained agent is used to check it's effectiveness on improving power system resilience against several windstorms. The performance of the agent for 9 unique line failure scenarios among these windstorms is given in Table II.

From Table II, it can be seen that the trained agent can maintain voltage stability with and without the trained agent violates for 2 and 6 scenarios, respectively. Thus, the proposed algorithm can enhance the resilience of the power system through controlling shunt reactive power output.

Also, from Table II, we can see that the trained agent cannot maintain the voltage stability for scenarios 9 and 10. This happens due to the fact that the shunts alone cannot maintain voltage magnitudes within limits.

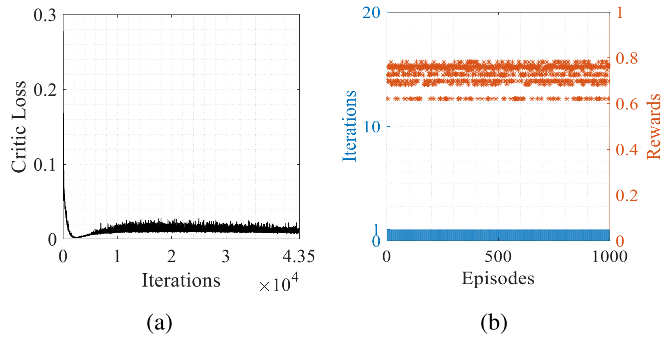


Fig. 3. (a) Losses of critics during training and (b) number of iterations and amount of rewards for testing episodes.

TABLE II. Resilience Enhancement Using the Trained Agent

S/L No.	Tripped Lines	Voltage Violations (Bus No.)	
		Without Agent	With Agent
1	2–5	None	None
2	1–3, 2–6, 2–5	5, 6, 7, 8	None
3	2–6, 2–5, 2–4	5, 6, 7, 8, 28	None
4	2–5, 2–4	7	None
5	2–5, 1–3, 2–4	7, 8	None
6	2–6, 2–5, 1–3, 2–4	None	None
7	2–4, 3–4, 2–5, 1–3	None	None
8	1–3, 2–5, 3–4, 2–6	All	All
9	2–6, 3–4, 2–5, 2–4, 1–3	All	All

V. Conclusion

This paper has proposed a data-driven-based approach for shunt dispatching to enhance voltage stability in power systems during windstorms, thus defining a new role for shunts in enhancing power system operational resilience. A multi-agent framework was developed using continuous soft actor critic algorithm to determine optimal shunt dispatch in case of single or multiple line outages. The algorithm was trained using hypothetical datasets, which were generated using system responses to outages and fragility curves of system components to windstorms. The trained algorithm was tested on several systems including the IEEE 30-bus system. The results showed that the proposed approach could maintain voltage magnitudes at system buses within the standard limits for most of the cases. Voltage magnitudes in some cases could not be maintained within the limits, which is not surprising. Since shunts alone cannot maintain voltage magnitudes within limits in some cases, the proposed approach can be extended (or integrated with existing algorithms) to include other resources such as generation dispatch, load shedding, and reconfiguration, to name a few.

Acknowledgement

This work was supported in part by the U.S. National Science Foundation (NSF) under Grant NSF 1847578.

References

[1] N. Bhusal, M. Abdelmalak, M. Kamruzzaman, and M. Benidris, "Power system resilience: Current practices, challenges, and future directions," *IEEE Access*, vol. 8, pp. 18 064–18 086, 2020.

[2] R. J. Campbell, "Weather-related power outages and electric system resiliency," Congressional Research Service, Tech. Rep., 2012.

[3] G. Huang, J. Wang, C. Chen, J. Qi, and C. Guo, "Integration of preventive and emergency responses for power grid resilience enhancement," *IEEE Trans. on Power Systems*, vol. 32, no. 6, pp. 4451–4463, 2017.

[4] H. Farzin, M. Fotuhi-Firuzabad, and M. Moeini-Aghtaie, "Enhancing power system resilience through hierarchical outage management in multi-microgrids," *IEEE Trans. on Smart Grid*, vol. 7, no. 6, pp. 2869–2879, 2016.

[5] M. Yan, X. Ai, M. Shahidehpour, Z. Li, J. Wen, S. Bahramira, and A. Paaso, "Enhancing the transmission grid resilience in ice storms by optimal coordination of power system schedule with pre-positioning and routing of mobile dc de-icing devices," *IEEE Trans. on Power Systems*, vol. 34, no. 4, pp. 2663–2674, 2019.

[6] X. Zeng, Z. Liu, and Q. Hui, "Energy equipartition stabilization and cascading resilience optimization for geospatially distributed cyber-physical network systems," *IEEE Trans. on Systems, Man, and Cybernetics: Systems*, vol. 45, no. 1, pp. 25–43, 2015.

[7] J. Li, F. Liu, Y. Chen, C. Shao, G. Wang, Y. Hou, and S. Mei, "Resilience control of dc shipboard power systems," *IEEE Trans. on Power Systems*, vol. 33, no. 6, pp. 6675–6685, 2018.

[8] Z. Li, M. Shahidehpour, F. Aminifar, A. Alabdulwahab, and Y. Al-Turki, "Networked microgrids for enhancing the power system resilience," *Proceedings of the IEEE*, vol. 105, no. 7, pp. 1289–1310, 2017.

[9] L. Sun, W. Liu, C. Y. Chung, M. Ding, R. Bi, and L. Wang, "Improving the restorability of bulk power systems with the implementation of a wf-bess system," *IEEE Trans. on Power Systems*, vol. 34, no. 3, pp. 2366–2377, 2019.

[10] C. Wang, Y. Hou, F. Qiu, S. Lei, and K. Liu, "Resilience enhancement with sequentially proactive operation strategies," *IEEE Trans. on Power Systems*, vol. 32, no. 4, pp. 2847–2857, 2017.

[11] A. S. Musleh, H. M. Khalid, S. M. Muyeen, and A. Al-Durra, "A prediction algorithm to enhance grid resilience toward cyber attacks in wamcs applications," *IEEE Systems Journal*, vol. 13, no. 1, pp. 710–719, 2019.

[12] M. Nazemi, M. Moeini-Aghtaie, M. Fotuhi-Firuzabad, and P. Dehghanian, "Energy storage planning for enhanced resilience of power distribution networks against earthquakes," *IEEE Trans. on Sustainable Energy*, vol. 11, no. 2, pp. 795–806, 2020.

[13] Y. Wang, L. Huang, M. Shahidehpour, L. L. Lai, H. Yuan, and F. Y. Xu, "Resilience-constrained hourly unit commitment in electricity grids," *IEEE Trans. on Power Systems*, vol. 33, no. 5, pp. 5604–5614, 2018.

[14] B. Taheri, A. Safdarian, M. Moeini-Aghtaie, and M. Lehtonen, "Enhancing resilience level of power distribution systems using proactive operational actions," *IEEE Access*, vol. 7, pp. 137 378–137 389, 2019.

[15] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor," *CoRR*, vol. abs/1801.01290, 2018. [Online]. Available: http://arxiv.org/abs/1801.01290

[16] Z. Fan, R. Su, W. Zhang, and Y. Yu, "Hybrid actor-critic reinforcement learning in parameterized action space," *CoRR*, vol. abs/1903.01344, 2019. [Online]. Available: http://arxiv.org/abs/1903.01344

[17] T. Haarnoja, A. Zhou, K. Hartikainen, G. Tucker, S. Ha, J. Tan, V. Kumar, H. Zhu, A. Gupta, P. Abbeel, and S. Levine, "Soft actor-critic algorithms and applications," *CoRR*, vol. abs/1812.05905, 2018.

[18] S. Fujimoto, H. van Hoof, and D. Meger, "Addressing function approximation error in actor-critic methods," *CoRR*, vol. abs/1802.09477, 2018.

[19] A. Hussain, A. Oulis Rousis, I. Konstantelos, G. Strbac, J. Jeon, and H. Kim, "Impact of uncertainties on resilient operation of microgrids: A data-driven approach," *IEEE Access*, vol. 7, no. DOI: 10.1109/ACCESS.2019.2891786, pp. 14 924–14 937, Jan. 2019.

[20] G. Li, P. Zhang, P. B. Luh, W. Li, Z. Bie, C. Serna, and Z. Zhao, "Risk analysis for distribution systems in the northeast u.s. under wind storms," *IEEE Trans. on Power Systems*, vol. 29, no. 2, pp. 889–898, 2014.

[21] A. Singhal, V. Ajjarapu, J. Fuller, and J. Hansen, "Real-time local volt/var control under external disturbances with high pv penetration," *IEEE Trans. on Smart Grid*, vol. 10, no. 4, pp. 3849–3859, 2019.

[22] G. Qu and N. Li, "Optimal distributed feedback voltage control under limited reactive power," *IEEE Trans. on Power Systems*, vol. 35, no. 1, pp. 315–331, 2020.

[23] S. Wang, J. Duan, D. Shi, C. Xu, H. Li, R. Diao, and Z. Wang, "A data-driven multi-agent autonomous voltage control framework using deep reinforcement learning," *IEEE Trans. on Power Systems*, vol. DOI: 10.1109/TPWRS.2020.2990179, pp. 1–11, 2020.

[24] R. Lincoln. Pypower. [Online]. Available: https://pypi.org/project/PYPOWER/