Aggregating Learning Agents for Microgrid Energy Scheduling During Extreme Weather Events

Avijit Das, Zhen Ni, and Xiangnan Zhong
Dept. of Computer, Electrical Engineering and Computer Science
Florida Atlantic University, Boca Raton, FL 33473
Email: {adas2017, zhenni, xzhong}@fau.edu

***Abstract*—Efficient utilization of the microgrid generation resources (MGRs) is important especially during weather-related events. Existing single-agent and cooperative multi-agent based reinforcement learning (RL) approaches may be infeasible and computationally expensive for scheduling MGRs effectively when there are some probabilistic weather-related emergency events. In this paper, we propose a Q-learning approach with multiple local agents for a grid-connected microgrid application and show that the proposed integration is capable of scheduling MGRs efficiently for both normal and weather-related emergency events. Specifically, we utilize different local Q-learning agents to learn different microgrid events, and aggregate the learned value functions to the global agent who handles the overall microgrid energy scheduling in a probabilistic way. Numerical simulations are performed to validate the effectiveness of the proposed method. The influences of the effective utilization of the MGRs and power outage duration are discussed. Two case studies with different power outage probabilities are presented to evaluate the performance of our proposed method.**

***Index Terms*—Microgrid energy scheduling, extreme weather events, energy optimization, reinforcement learning, and aggregating knowledge.**

I. Introduction

Increasing threats of weather-related incidents and natural disasters in recent years have highlighted the urgency of effectively improving power system resilience and attracted worldwide attention. In the U.S., between the years 2003-2012, more than 10 million, and after 2012-till now, higher than 17 million customers experienced a power outage due to weather-related events [1], [2], [3]. The annual impact of weather-related blackouts costs between \$20 to \$75 billion only in the U.S., and the cost keeps increasing every year [4], [5]. Deployment of microgrids for improving power system resiliency is the widely accepted viable solution. With the ability to operate in both grid-connected and isolated mode, microgrid enhances reliability and resiliency, increases efficiency, and provides cleaner and cheaper energy through enabling a diverse distributed energy mix [6], [7]. Therefore, efficient pre-disaster preparation and proper utilization of microgrid energy sources can significantly minimize the economic loss and power interruptions.

As stated in [8], microgrid scheduling and dispatch research with resiliency considerations have not been widely explored in the literature. A probabilistic chance constraint is proposed in order to meet the local microgrid demands considering renewable generation (RG) and demand uncertainty in [9]. A new optimization model is presented for determining the spinning reserve requirement in microgrids by analyzing the characteristic of unit outage events in [10]. A microgrid scheduling and dispatch model is proposed and evaluated considering the main grid supply interruption time and duration in [11]. This work was later extended considering RG and load uncertainties in [12]. However, in these works, the authors have assumed the energy storage charging/discharging status to be the same for both normal and isolated operations that may output an extra-operating cost in practice. Some works study the natural disaster impacts, trying to understand the blackout causes and explore ways to prepare the grid [13], [14].

In recent years, single- and multi-agent reinforcement learning (RL) approaches have been recognized for solving microgrid online decision-making and control problems [15]. In [5], a multi-agent reinforcement learning (RL) technique is proposed to improve the microgrid post-disaster resilience with the goal to minimize outage duration. Though post-disaster resilience can help to minimize the outage duration through scheduling available spinning reserves, pre-disaster planning and efficient utilization during a disaster can be a cost-effective solution with significant enhancement of flexibility, reliability, and resiliency. Multi-agent RL with cooperative system is explored in the literature for solving microgrid online energy management challenges [16]. In cooperative RL, the agents cooperatively interact with the environment, and the number of states increases significantly with the increment of sub-environments so that this approach computationally is sometimes expensive. Asynchronous and synchronous RL approaches are proposed in the literature [17], [18]. In these approaches, the local agents are employed in parallel to explore different parts of the environment and update the learned policy asynchronously or synchronously to the global network to handle complex tasks.

Under the light of the new artificial intelligence trend, in this paper, we propose a Q-learning approach with multiple agents to capture different microgrid operating conditions when the grid is suffering from extreme weather. For developing our proposed approach, we utilize the parallel learning nature so that the proposed approach is capable of solving both single (single task) and multiple (multi-task) microgrid environment problems efficiently. We apply the proposed approach for a grid-connected microgrid application with a probabilistic Q-value updating strategy. In the proposed design, local Q-learning agents are used to learn the optimization policy for normal and weather-related emergency events, respectively. The global Q-learning agent updates the Q-value function probabilistically based on each local agent's learned policy, and thus takes the proper microgrid dispatch decision based on the corresponding events. The energy scheduling performance of the proposed algorithm is evaluated through numerical

2021 IEEE Power & Energy Society General Meeting (PESGM) | 978-1-6654-0507-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/PESGM46819.2021.9637949

case studies and compared with the existing techniques (e.g., cooperative Q learning) to provide the improvement.

The rest of this paper is organized as follows. The model description and problem formulation are presented in Section II. In Section III, the integration of the proposed Q-learning approach in microgrid application is demonstrated. Simulation results and analysis are carried out in Section IV. Finally, the conclusions and future works are presented in Section V.

II. Model Description and Problem Formulation

A microgrid can be defined as a small distribution system consisting of distributed energy resources (DERs), including distributed RGs, dispatchable distributed generators, and battery energy storage system (BESS) with the interconnection of controllable and uncontrollable loads [19], [20]. In this paper, we consider a grid-connected microgrid with four types of units from the standpoint of energy generation and demand: intermittent RGs, such as wind turbines and photovoltaics; a diesel generator (DG) as a dispatchable distributed generator; a BESS; and local microgrid loads. The connection with the main grid gives the flexibility to the microgrid to export/import power to/from the utility network, and maintain the reference voltage and frequency of the system in accordance with the predetermined operation strategy. In this problem, our goal is to schedule the generation units efficiently based on the defined probabilistic weather conditions so that the total operational cost of the microgrid can be minimized. We solve the optimization problem for a day with an hour interval. The problem is formulated as a Markov decision process, where the state variables represent the microgrid input information, and the action set represents the microgrid scheduling decision variables. The microgrid state is defined as

$$S_t = (B_t, R_t, G_t, D_t). \tag{1}$$

where t is the time step, B_t is the available energy in the BESS, R_t is the available RG output, G_t is the grid price, and D_t is the microgrid load demand. The action policy set is

$$a_t = (a_t^{B,c}, a_t^{B,d}, a_t^{DG}, a_t^{m,G}, a_t^{G,m}, a_t^{dl}), a_t \in \chi_t. \tag{2}$$

where $a_t^{B,c}$ and $a_t^{B,d}$ represent charging and discharging power of the BESS, respectively. a_t^{DG} is the DG power output. $a_t^{m,G}$ and $a_t^{G,m}$ represent the export and import powers to and from the main grid, respectively. a_t^{dl} is the dumped or unserved load.

The microgrid instant cost function is defined as the summation of the cost of energy buying from and selling to the grid and the fuel cost of the dispatchable DG unit as

$$C(S_t, a_t) = (a_t^{G,m} - a_t^{m,G})G_t + k_{DG,t}(x(a_t^{DG})^2 + y a_t^{DG} + z). \tag{3}$$

where x, y and z are the DG fuel cost-curve coefficients. $k_{DG,t}$ is a binary variable acting the ON/OFF status of the DG.

The microgrid operational constraints are as follows

$$a_t^{DG} + a_t^{G,m} + a_t^{B,d} - a_t^{B,c} - a_t^{m,G} + a_t^{dl} + R_t = D_t, \tag{4}$$

$$0 \le a_t^{B,c} \le (1 - b_t)\psi^C, \tag{5}$$

$$0 \le a_t^{B,d} \le b_t \psi^D, \tag{6}$$

$$SOC_{\min} \le SOC_t \le SOC_{\max}, \tag{7}$$

$$k_{\text{gen}} P_{\text{rated}} k_{DG,t} \le a_t^{DG} \le P_{\text{rated}} k_{DG,t}. \tag{8}$$

where, the constraint (4) is the microgrid generation-demand balance constraint. The constraints (5) and (6) are the BESS charging and discharging power output constraints where ψ^C and ψ^D represent the maximum charging and discharging battery power output, respectively. The BESS state of charge (SOC) constraint is presented in (7) to keep the SOC within a certain range for the healthy operation of the BESS. The DG power output should be in a certain range, and it is constrained using (8) where k_{gen} is defined as a percentage of the DG rated power P_{rated}.

A transition function is used to update the SOC of the BESS as

$$SOC_{t+1} = \frac{1}{B_{\text{cap}}}(B_{\text{cap}} SOC_t + \phi^C a_t^{B,c} - \frac{a_t^{B,d}}{\phi^D}). \tag{9}$$

where ϕ^C and ϕ^D are the BESS charging and discharging efficiency. B_{cap} is the energy capacity of the BESS.

The objective is to minimize the total operational cost of the microgrid over a finite horizon of time T,

$$\min_{a_t} \mathbb{E}\left[\sum_{t=0}^{T} C(S_t, a_t)\right]. \tag{10}$$

where $\mathbb{E}[.]$ is the expectation operator. The objective function subjects to the microgrid operational constraints presented in (4)-(8).

III. Proposed Q-learning Approach with Multiple Local Agents for the Microgrid Application

In RL research, the RL agent interacts with the environment (system model) through taking actions/decisions online and learning the policy from the local observation [15], [21]. Q-learning is an RL algorithm, and in Q-learning, a Q-value function $Q(s,a)$ is used to map the relationship between state s and action a. The Q-value function is calculated by Bellman equation as

$$Q(S_t, a_t) = (1-\alpha)Q(S_t, a_t) + \alpha[C(S_t, a_t) + \gamma \min_{a_{t+1}} Q(S_{t+1}, a_{t+1})] \tag{11}$$

where α is the learning rate, and S_{t+1} is the resulting state after taking action a_t in state S_t. In state S_t, the action a_t is usually determined using the ϵ-greedy strategy. According to the strategy, at any state S_t, the a_t is determined either selecting a random action from the feasible actions or using the greedy technique $\min_{a_t} Q(S_t, a_t)$. γ is the discount factor.

In this paper, we propose a Q-learning approach with multiple local agents to solve the microgrid scheduling problem under the resiliency considerations. Specifically, we employ two local Q-learning agents to learn the optimization policies parallelly for both normal and weather-related emergency operations of the grid-connected microgrid. Both agents are probabilistically connected to the global Q-agent so that the global microgrid policy can be updated, and the microgrid dispatch decisions can be taken based on event probability.

The integration of the proposed Q-learning approach in the microgrid application is illustrated in Fig. 1. According to the algorithm design, Q-table and other microgrid parameters are initialized at the beginning of the algorithm. The local agents interact with the sub-environments (normal and emergency) of

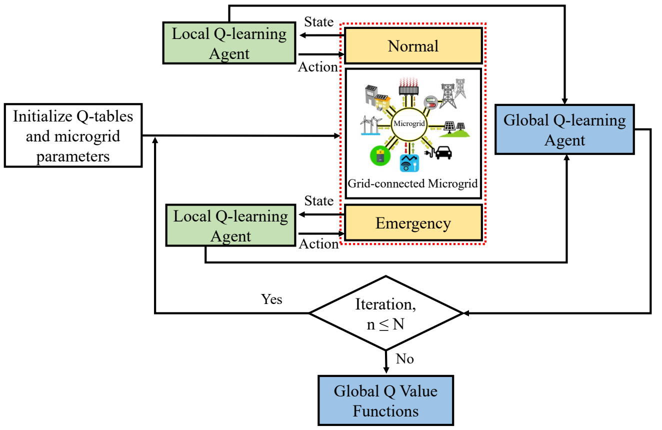

Fig. 1. Integration of the proposed Q-learning approach in the grid-connected microgrid application.

the micorgrid environment and learn the output of the decision taken by updating the Q-table of the corresponding state-action pair. The actions of the local agents are taken using the ϵ-greedy strategy where the greedy decisions of the local agents for the normal and emergency operations are defined as

$$a_t^{nr} = \arg\min_{a_t^{nr}} Q^{nr}(S_t^{nr}, a_t^{nr}). \tag{12}$$

$$a_t^{emg} = \arg\min_{a_t^{emg}} Q^{emg}(S_t^{emg}, a_t^{emg}). \tag{13}$$

where Q^{nr} and Q^{emg} are the Q-value value functions for the normal operation and emergency operation agents, respectively. S_t^{nr} and S_t^{emg} represent the states obtained from the normal and emergency sub-environments of the microgrid model. Basically, these states represent the available microgrid resource information under different microgrid event conditions.

At every iteration, the Q-value functions of the global Q-learning agent at time t is updated probabilistically as

$$Q^G(s_t, a_t) = (1-p)Q^{nr}(s_t, a_t) + pQ^{emg}(s_t, a_t). \tag{14}$$

where Q^G is the Q-value function for the global agent. p represents the event probability for the emergency operation. s_t represents the set of states at time t, and a_t is the set of actions per state. This equation provides the global agent the capacity to aggregate the knowledge learned from both the normal and the emergency operation conditions. Thus, the global agent is able to handle the extreme weather events online adaptively.

At the end of N iterations, the proposed approach outputs the global Q-value functions that can be used to determine the microgrid power dispatch decisions as the greedy policy $\arg\min_{a_t} Q^G(S_t, a_t)$ for the corresponding probabilistic events. Besides, the microgrid operator can also access the trained local Q^{nr} and Q^{emg} value functions that give flexibility to the operator to decide the microgrid operational strategy based on real-time events.

IV. Simulation Results and Analysis

In this section, we conduct different case studies to examine the performance of the proposed approach in terms of microgrid operational cost and final action policy. We also present the performance comparison to justify the performance improvement.

A. Simulation Setup

The parameters of the microgrid DERs are provided in Table I. The time horizon of of the optimization problem is set to be $T = 24$ hours with one hour interval. The microgrid input profiles including a small residential community load-demand, RG output, and electricity price are plotted in Fig. 2. The RG outputs are taken from the system advisory model by National Renewable Energy Laboratory for the city of Phoenix, AZ [22]. For the load-demand, a small residential community load-demand data is collected from [23].

TABLE I
Microgrid information.

RG	Photovoltaic Capacity	50 kW
	Wind Turbine Capacity	50 kW
BESS	Capacity	150 kWh
	charging and discharging eff.	90%
	Maximum Power	30 kW
DG	Rated Power	100 kW
	Minimum Dispatch Percentage	0.3
	Cost Coefficients x, y, and z	0.0009 (\$/$(kW)^2$), 0.0213 (\$/kW) and 1.1 (\$)

For the implementation of the Q-learning approach, we use the lookup tables for the agents Q^G, Q^{nr} and Q^{emg}. We set the lookup tables as a matrix of states and time steps, where discretized battery SOC and DG ON/OFF status ($k_{DG,t}$) are used for defining the state. We use the discretized states to avoid the computational burden of solving problems with continuous states. The exploration rate ϵ is set as 0.6. For investigating the impact of power supply interruption on microgrid operation, we assume during the extreme weather-related event, the utility grid goes down for a certain time frame, and the microgrid operates in isolated mode on that certain time frame. According to the U.S. Energy Information Administration (EIA) [24], U.S. customers experienced an average of around 6 hours of power interruptions in 2018 including major and non-major events, where around 4 hours of power interruptions only for the major events. In our case study, we use 4 hours of power interruptions of the utility grid on the microgrid operation.

All the simulations are conducted in MATLAB R2019b environment on a 8th generation Intel Core i7 8650U quad-core processor, 4.2GHz Windows based PC with 16GB RAM. For the performance comparison, all the approaches are implemented in the same environment.

B. Simulation Results

In this section, we conduct two case studies with different probability rates of weather-related events and also consider the impact of utility power interruptions on microgrid operation for a certain time frame.

1) Case Study 1 - High Probability of Outage: In this case study, we assume the extreme weather-related event probability as 70%, and utility power interruptions are considered at hours $4-7$. Our proposed Q-learning approach interacts

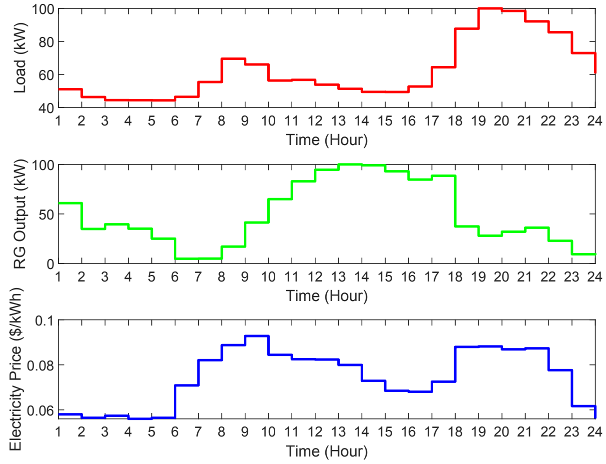

Fig. 2. Microgrid exogenous information, load, RG output, and electricity price from the utility grid.

with the microgrid environments online considering the given assumptions and outputs the microgrid optimization policy or scheduling decisions. We also apply the similar assumptions for other existing approaches, and the performance comparison is presented in Table II.

TABLE II
COMPARATIVE STUDY IN TERMS OF EXPECTED MICROGRID TOTAL OPERATIONAL COST FOR CASE STUDY 1.

Approach	DP (Offline)	Proposed Q-learning	Cooperative Q-learning
Expected Total Operation Cost ($)	20.4	21.2	23.3

We obtain the optimal solution using the dynamic programming (DP) approach. Note, the DP approach is an offline optimization technique, and this technique can not be used online for determining microgrid scheduling decisions. We use 4000 iterations for the Q-learning approaches, and the average results are presented after 50 runs of simulations. In Table II, the results show that the proposed Q-learning approach can obtain the expected total microgrid operational cost as $21.2 that is very close to the reference value with only a $0.8 gap. Note, this gap reduces with the increment of the iteration number. The multi-agent cooperative Q-learning approach obtains the expected total operational cost as $23.3 with a $2.9 gap. It indicates that this approach requires more intensive training with additional computation costs. The comparative study shows that our proposed approach outperforms the existing cooperative Q-learning approach.

The convergence curves in terms of microgrid average total operational cost are plotted in Fig. 3(a). These cost curves are representing the expected total cost over the number of iterations. The results show that the cost curve of the proposed Q-learning approach drops faster than the cooperative Q-learning approach and converges close to the optimal solution with a lower number of iterations. The power outputs of the microgrid resources obtained from the proposed approach are also plotted in Fig. 3(b). During the utility power interruption time-period, the proposed approach efficiently utilizes the DG and BESS units to keep microgrid operation uninterrupted.

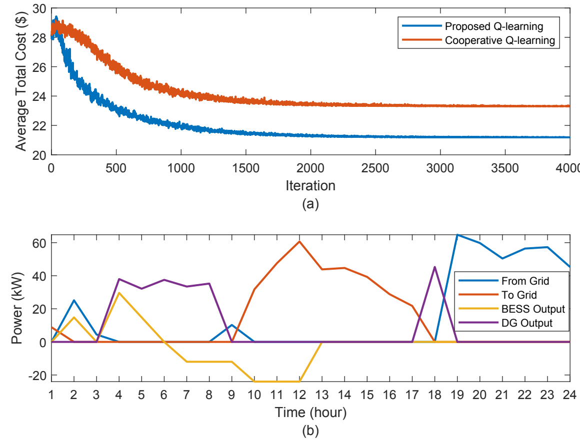

Fig. 3. Microgrid average total cost curve convergence study and the power outputs of the microgrid resources for case study 1. At any time t, the BESS output is defined as $a_t^{B,c} - a_t^{B,d}$, so the positive value represents the charging power, and the negative value represents the discharging power.

According to the result analysis, the proposed Q-learning approach can efficiently schedule microgrid energy resources during weather-related emergency events and minimize the operation cost.

2) Case Study 2 - Low Probability of Outage: In this case study, we assume that the extreme weather-related event hits during the time frame $4-7$, and the probability of utility power interruptions is 30%. The expected microgrid operation costs obtained from different approaches are presented in Table III.

TABLE III
COMPARATIVE STUDY IN TERMS OF EXPECTED MICROGRID TOTAL OPERATIONAL COST FOR CASE STUDY 2.

Approach	DP (Offline)	Proposed Q-learning	Cooperative Q-learning
Expected Total Operation Cost ($)	19.4	20.1	22.3

The results show that the expected average total cost of the microgrid obtained from the proposed online Q-learning approach is $20.1. With a similar simulation setup, the cooperative Q-learning outputs the expected cost as 22.3. The comparative study in terms of expected cost curve convergence is illustrated in Fig. 4(a). In the figure, we can observe that the cost curve of the proposed Q-learning approach rapidly drops over the iteration and converges close to the reference value with a $0.7 gap. On the other hand, the existing cooperative Q-learning approach gradually drops towards the optimal value and converges at $22.3 with a $2.9 gap from the reference DP approach. The global optimization policy of the proposed Q-learning approach for the corresponding expected cost is illustrated in Fig. 4(b). Since the outage probability is 30%, the obtained microgrid optimization policy is highly influenced by the microgrid normal operation. The microgrid schedules power from the grid to charge the battery and fulfill the surplus load-demands during the time frame $4-7$ hours considering the future outcome of the current taken actions.

The proposed Q-learning approach has a unique advantage over the existing Q-learning approaches. In the proposed inte-

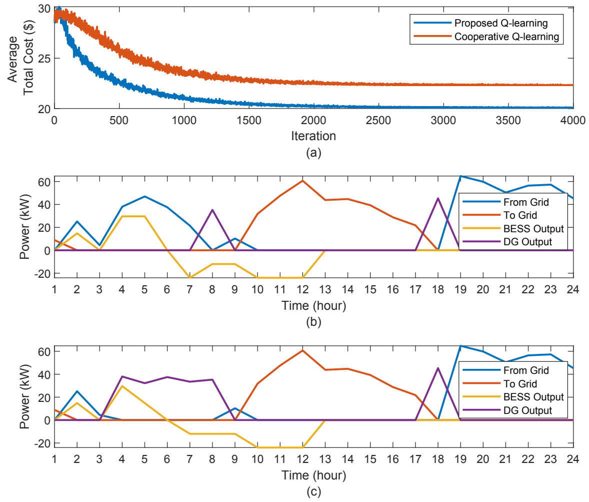

Fig. 4. Microgrid expected average total cost curve convergence study and the power outputs of the microgrid resources for case study 2. Here, the plot (b) presents the power outputs obtained from the global agent, and the plot (c) presents the power outputs obtained from the Q^{emg} agent.

gration, the microgrid operator can also obtain the optimization policy for a specific event using the greedy policy of the local Q-value functions. Even with low outage probability, the microgrid can obtain the emergency microgrid dispatch decisions and schedule the resources based on its operation strategy. The microgrid scheduling decisions for the emergency operation obtained using the local Q^{emg}-value functions are presented in Fig. 4(c). The results show that the obtained microgrid decisions are similar to the decisions presented in Fig. 3(b). It indicates that the proposed integration provides flexibility to the microgrid operators to generate the emergency microgrid scheduling decisions even with low outage probability. Therefore, the proposed Q-learning approach can be a promising tool for solving microgrid scheduling problems with resiliency considerations.

V. Conclusion

In this paper, we propose a Q-learning approach with multiple local learning agents to handle the weather-related emergency events in a grid-connected microgrid. In our proposed algorithm design, different local agents are used to learn different microgrid events. The learned value functions are aggregated to update the global agent's policy in a probabilistic way. Numerical simulations are performed to validate the effectiveness of the proposed method for a certain power outage duration. The results show that the proposed Q-learning approach can efficiently schedule the microgrid DERs, considering the probabilistic emergency operation, and minimize the operational cost. An interesting future work would be making the proposed approach capable of handling the randomness of the RGs and apply the proposed algorithm for solving real-time decision-making problems with uncertainty.

Acknowledgment

This work is partly supported by the National Science Foundation under grant 1949921 and 1947418. This work is also partly supported by ORAU Ralph E. Powe Junior Faculty Enhancement Awards.

References

[1] F. H. Jufri, V. Widiputra, and J. Jung, "State-of-the-art review on power grid resilience to extreme weather events: Definitions, frameworks, quantitative assessment methodologies, and enhancement strategies," *Applied Energy*, vol. 239, pp. 1049–1065, 2019.

[2] U. S. DOE, *Hurricane Matthew Situation Reports October 2016*, https://www.energy.gov/oe/downloads/hurricane-matthew-situation-reports-october-2016.

[3] ——, *Hurricane Michael Situation Reports - October 2018*, https://www.energy.gov/ceser/downloads/hurricane-michael-situation-reports-october-2018.

[4] M. Panteli and P. Mancarella, "Influence of extreme weather and climate change on the resilience of power systems: Impacts and possible mitigation strategies," *Electric Power Systems Research*, vol. 127, pp. 259–270, 2015.

[5] H. Nie, Y. Chen, Y. Xia, S. Huang, and B. Liu, "Optimizing the post-disaster control of islanded microgrid: A multi-agent deep reinforcement learning approach," *IEEE Access*, vol. 8, pp. 153 455–153 469, 2020.

[6] U. S. DOE, *Smart Grid System Report*, November, 2018, https://www.energy.gov/oe/downloads/2018-smart-grid-system-report.

[7] A. Das and Z. Ni, "A computationally efficient optimization approach for battery systems in islanded microgrid," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 6489–6499, 2017.

[8] G. Liu, T. B. Ollis, Y. Zhang, T. Jiang, and K. Tomsovic, "Robust microgrid scheduling with resiliency considerations," *IEEE Access*, vol. 8, pp. 153 169–153 182, 2020.

[9] B. Zhao, Y. Shi, X. Dong, W. Luan, and J. Bornemann, "Short-term operation scheduling in renewable-powered microgrids: A duality-based approach," *IEEE Transactions on sustainable energy*, vol. 5, no. 1, pp. 209–217, 2013.

[10] M. Q. Wang and H. Gooi, "Spinning reserve estimation in microgrids," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1164–1174, 2011.

[11] A. Khodaei, "Microgrid optimal scheduling with multi-period islanding constraints," *IEEE Transactions on Power Systems*, vol. 29, no. 3, pp. 1383–1392, 2013.

[12] ——, "Resiliency-oriented microgrid optimal scheduling," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 1584–1591, 2014.

[13] R. Nateghi, S. D. Guikema, and S. M. Quiring, "Comparison and validation of statistical methods for predicting power outage durations in the event of hurricanes," *Risk Analysis: An International Journal*, vol. 31, no. 12, pp. 1897–1906, 2011.

[14] Y. Wang, C. Chen, J. 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, 2015.

[15] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[16] F. Luo, Y. Chen, Z. Xu, G. Liang, Y. Zheng, and J. Qiu, "Multiagent-based cooperative control framework for microgrids' energy imbalance," *IEEE transactions on industrial informatics*, vol. 13, no. 3, pp. 1046–1056, 2016.

[17] J. X. Wang, Z. Kurth-Nelson, D. Tirumala, H. Soyer, J. Z. Leibo, R. Munos, C. Blundell, D. Kumaran, and M. Botvinick, "Learning to reinforcement learn," *arXiv preprint arXiv:1611.05763*, 2016.

[18] V. Mnih, A. P. Badia, M. Mirza, A. Graves, T. Lillicrap, T. Harley, D. Silver, and K. Kavukcuoglu, "Asynchronous methods for deep reinforcement learning," in *International conference on machine learning*, 2016, pp. 1928–1937.

[19] D. Wu, X. Ma, S. Huang, T. Fu, and P. Balducci, "Stochastic optimal sizing of distributed energy resources for a cost-effective and resilient microgrid," *Energy*, p. 117284, 2020.

[20] A. Das and Z. Ni, "A novel fitted rolling horizon control approach for real-time policy making in microgrid," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3535–3544, 2020.

[21] Z. Ni, S. Paul, X. Zhong, and Q. Wei, "A reinforcement learning approach for sequential decision-making process of attacks in smart grid," in *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, 2017, pp. 1–8.

[22] N. Blair, A. P. Dobos, J. Freeman, T. Neises, M. Wagner, T. Ferguson, P. Gilman, and S. Janzou, "System advisor model, sam 2014.1. 14: General description," National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep. NREL/TP-6A20-61019, Feb. 2014.

[23] NREL, *openei*, https://openei.org/community/blog/commercial-and-residential-hourly-load-data-now-available-openei.

[24] U. S. EIA, *Annual Electric Power Industry Report*, June 2020, https://www.eia.gov/electricity/data/eia861/.