Optimal Coordination of Distributed Energy Resources Using Deep Deterministic Policy Gradient

Avijit Das and Di Wu
Pacific Northwest National Laboratory
Contact: avijit.das@pnnl.gov

July 20, 2023

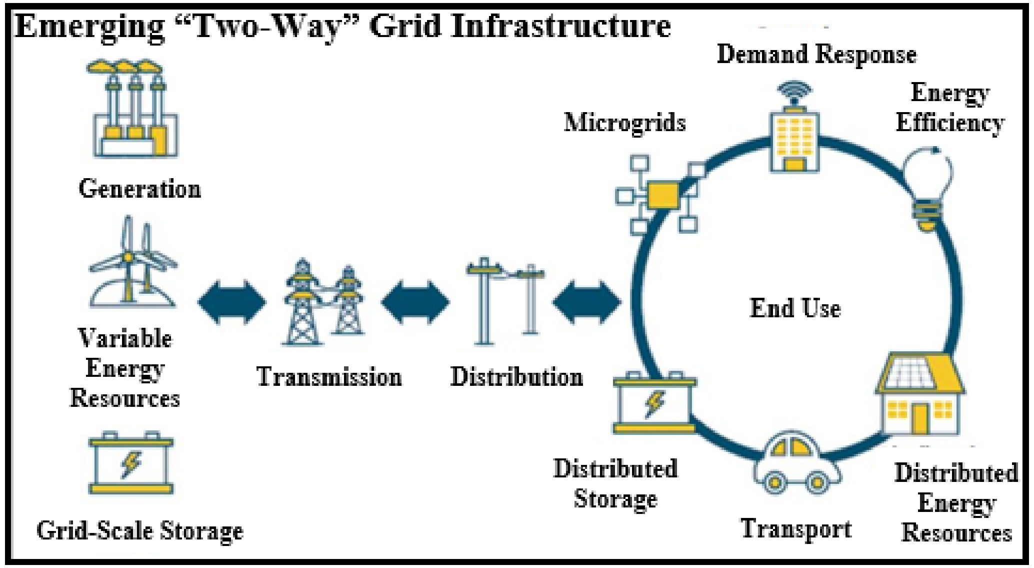
This material is based on work supported by the U.S. Department of Energy (DOE), Office of Electricity through Energy Storage Program. Pacific Northwest National Laboratory is operated for the DOE by Battelle Memorial Institute under Contract DE-AC05-76RL01830.

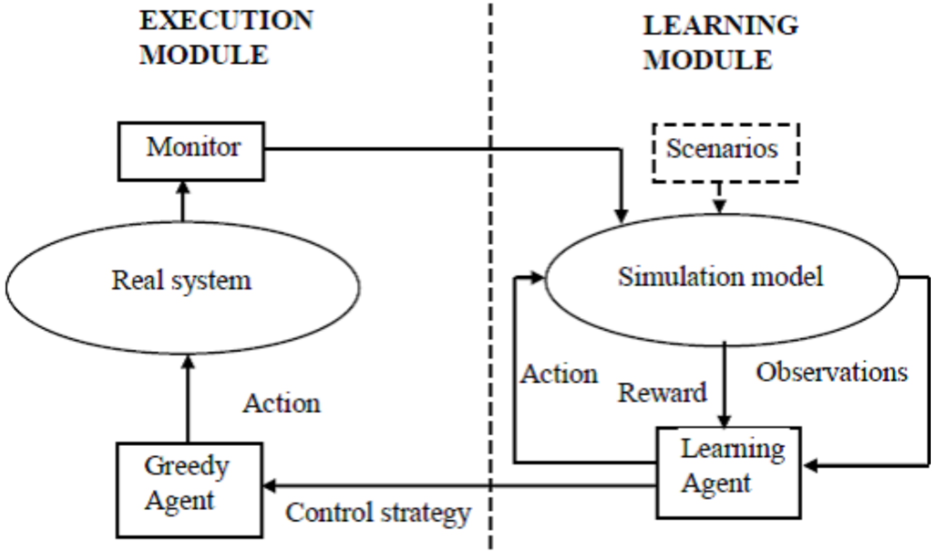
Outline

- Background
- Motivation and Contribution
- Microgrid System Model and Formulation
- Proposed Deep Deterministic Policy Gradient Approach
- Simulation Analysis
- Conclusion and Future Work

Background

- Distributed energy resource (DER) deployment in distribution systems has increased considerably
- Coordination of DERs amidst uncertainties is critical to harvesting their potential benefits
- Reinforcement learning (RL) is promising for power system applications, and has attracted surging attention due to its recent successes





Existing Methods

- Various RL methods have been proposed and investigated in recent years to tackle DER coordination challenges
  - Fitted-Q iteration algorithm to maximize self-consumption of renewable generation (RG) and minimize the electricity cost
  - A multiagent RL framework empowers autonomous agents of DERs and consumers for maximizing individual profit
  - Deep Q-learning method for coordination of battery energy storage systems (BESSs) considering microgrid system uncertainties
  - Deep RL approach to manage the optimal energy consumption of multiple smart homes with RG and BESS
  - Approximate dynamic programming with policy-based exploration for DER coordination in a remote microgrid
  - Deep deterministic policy gradient (DDPG) for optimal scheduling of DERs and service restoration

Motivation

- Shortcomings of existing methods:
  - Many of them are based on lookup table methods, which become inefficient when the problem size is large and infeasible when continuous states and actions are involved
  - BESS loss of life and degradation are not reasonably considered in the RL design

Contribution

- We propose an innovative DDPG-based RL for optimal DER dispatch with BESS loss of life explicitly modeled
  - Both calendrical and cyclical aging effects are taken into account when designing the dispatch policy

* A. Das and D. Wu. "Optimal Coordination of Distributed Energy Resources Using Deep Deterministic Policy Gradient." in *Proc. IEEE Power Energy Soc. Electr. Energy Stor. App. and Tech. Conf. (EESAT)*, Austin, TX, Nov. 2022, pp. 1-5.

Microgrid System Model

- Goal is to optimally dispatch DERs to minimize the cumulative operation cost
- DER includes
  - RG such as photovoltaic (PV) and wind
  - BESS assets
- Objective function consists of two components
  - BESS operation and maintenance (O&M) cost
  - Energy cost
- Component- and system-level constraints include
  - BESS power limit
  - BESS state of charge (SOC) transition and limit
  - Microgrid power balance

$$\min_{x} \sum_{t=1}^{T} (C_t^{\text{batt}} + C_t^{\text{ex}})$$

BESS O&M Cost

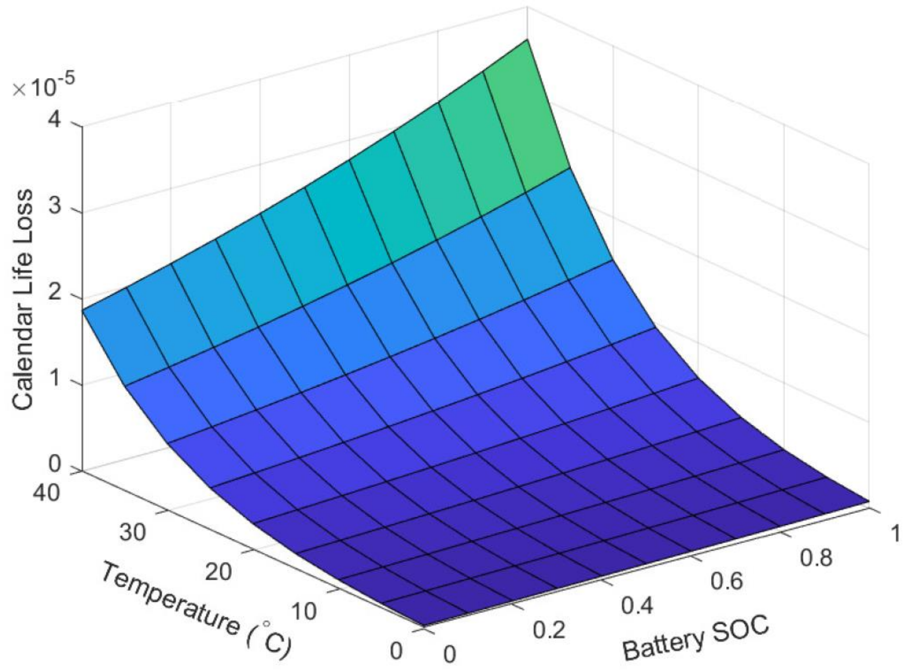
$$C_t^{\text{batt}} = \begin{cases} g p_t^{\text{batt}} \Delta t, if\, p_t^{\text{batt}} \geq 0 \; (discharging) \\ 0, \quad otherwise \; (charging) \end{cases}$$

Energy Cost

$$C_t^{\text{ex}} = \begin{cases} \dfrac{p_t^{\text{grid}} \lambda_t \Delta t}{\zeta}, if\, p_t^{\text{grid}} \geq 0 \\ p_t^{\text{grid}} \lambda_t \zeta \Delta t, if\, p_t^{\text{grid}} < 0 \end{cases}$$

BESS Life Loss

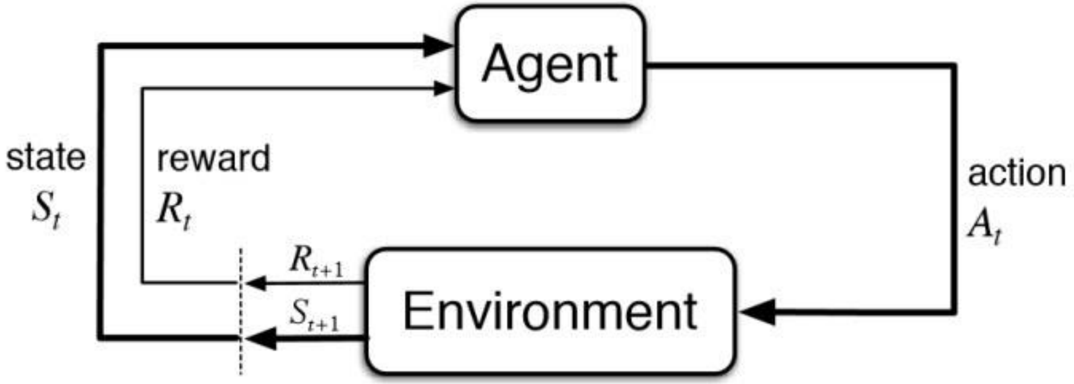
- This paper considers both cyclical and calendrical aging effects
- Cycle life loss is determined based on charging/discharging operation of BESS
  - Rainflow algorithm for determining the depth of discharge (DOD)
  - A cyclical aging model uses DOD to determine cycle life loss
- Calendar life loss is determined using a weighted modeling approach
  - Calendar life loss calculated due to staying at a certain SOC and temperature
  - Life loss is particularly prominent at high SOC
- Cumulative daily cycle life loss and calendar life loss are constrained to ensure the expect battery lifetime



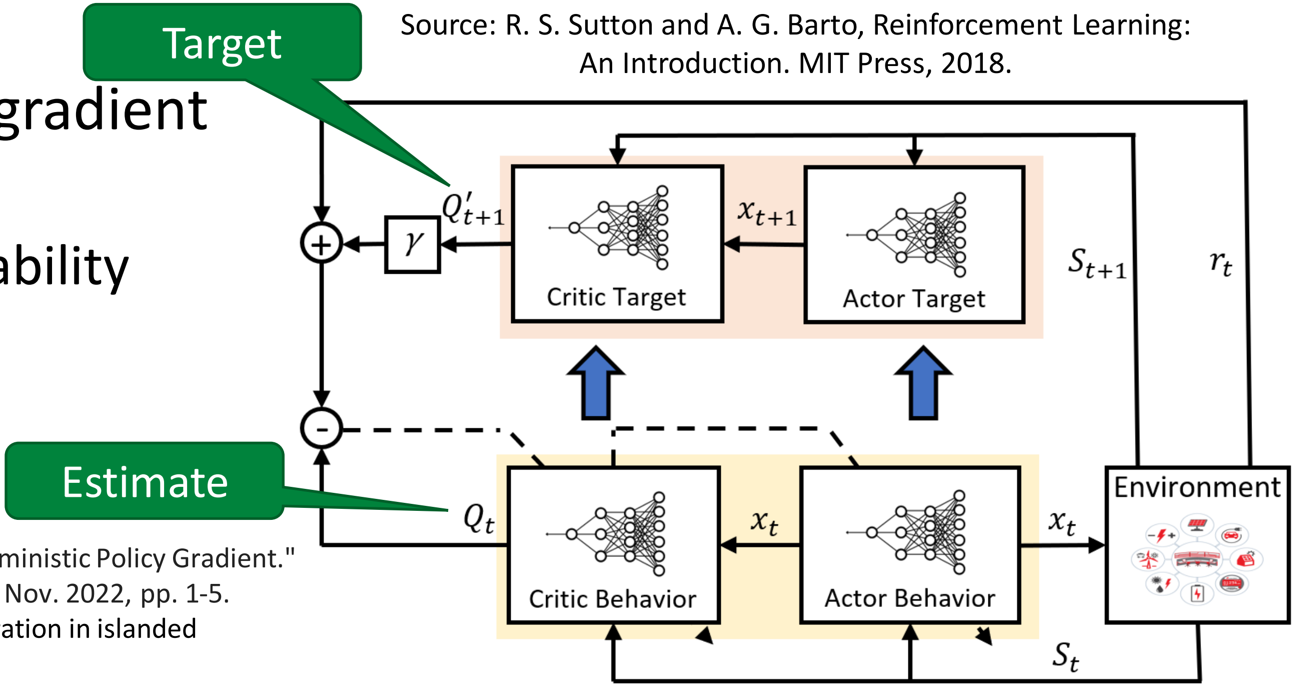
* A. Das and D. Wu. "Optimal Coordination of Distributed Energy Resources Using Deep Deterministic Policy Gradient." in *Proc. IEEE Power Energy Soc. Electr. Energy Stor. App. and Tech. Conf. (EESAT)*, Austin, TX, Nov. 2022, pp. 1-5.

Problem Formulation and DDPG

- Sequential decision-making problem is formalized as a Markov decision process
  - State includes battery SOC, RG output, system load, retail price, and cumulative life losses
  - Action includes BESS charging/discharging power, and the power purchased from/sold back to the grid
- DDPG is an off-policy deterministic policy gradient method
  - An off-policy method improves learning stability and data usage efficiency
  - Safe exploration is ensured through action mapping, clipping, and reward shaping



Source: R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction. MIT Press, 2018.

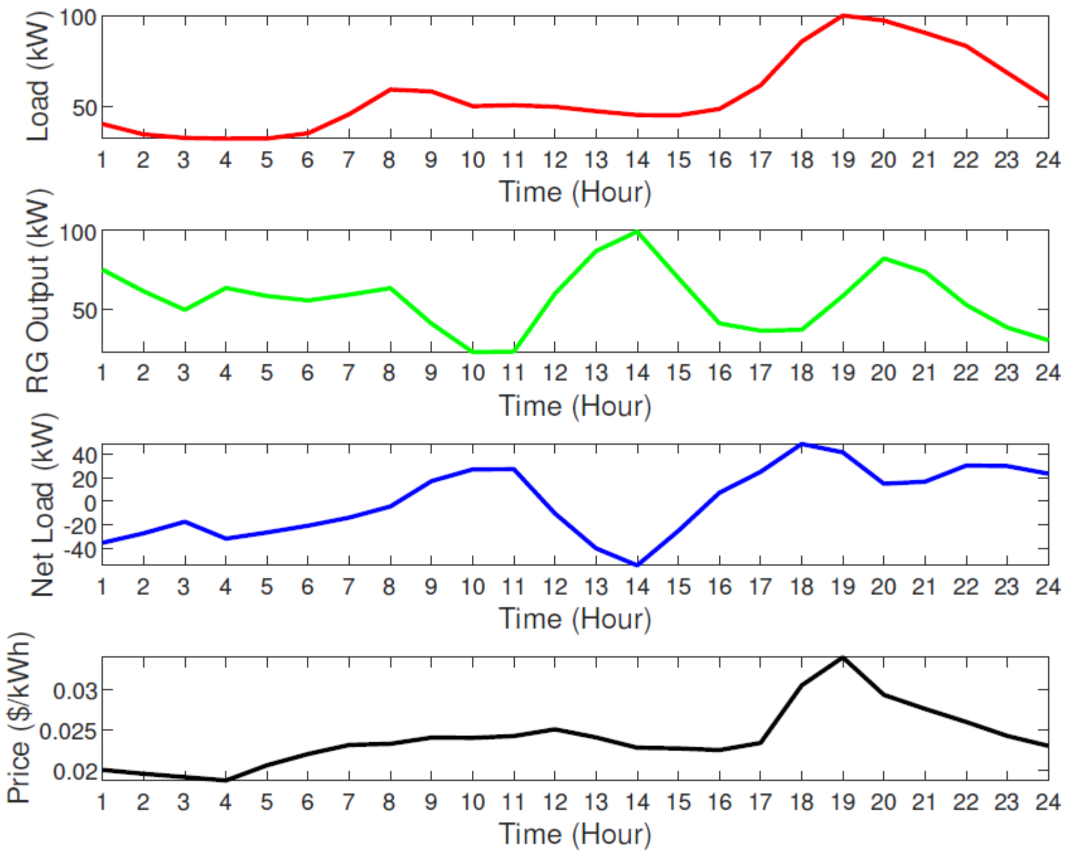


* A. Das and D. Wu. "Optimal Coordination of Distributed Energy Resources Using Deep Deterministic Policy Gradient." in *Proc. IEEE Power Energy Soc. Electr. Energy Stor. App. and Tech. Conf. (EESAT)*, Austin, TX, Nov. 2022, pp. 1-5.

* Y. Du and D. Wu, "Deep reinforcement learning from demonstrations to assist service restoration in islanded microgrids," *IEEE Trans. Sustain. Energy*, vol. 13, no. 2, pp. 1062–1072, Apr. 2022.

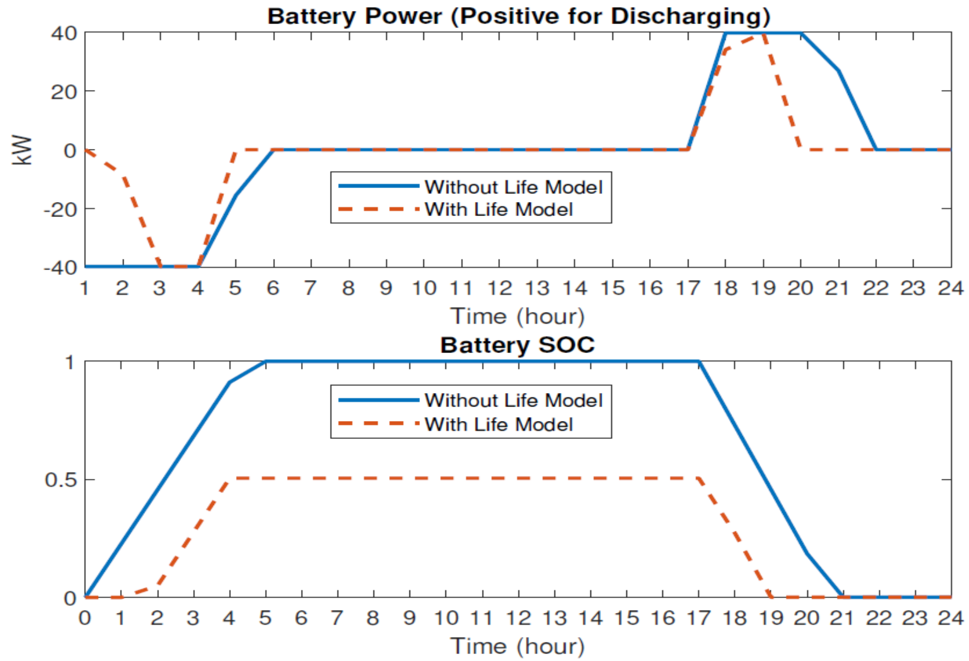
Simulation Setup

- Microgrid test system:
  - 100 kW wind and 50 kW PV
  - Lithium-ion BESS: 3 years of calendar life and 600 cycles
  - 160 kWh BESS energy capacity, 40 kW rated power, 83.7% round-trip efficiency
  - BESS SOC range: [0 1]
  - Residential load profiles with peak of 100 kW
- DDPG implemented using PyTorch
  - OpenAI gym for environment
  - Neural networks with two hidden layers and 64 neurons
  - ReLU and sigmoid activation functions



Case Study

- BESS dispatch results without the life loss model compared with the life loss model
  - BESS deep cycled when the life loss is not considered in control design, leading to reduced lifespan
  - Incorporating life-loss model into control design helps avoid deep cycling of BESS and expand its service life
- Different daily maximum life loss limits are also investigated
  - BESS expected lifetime is minimum when the life loss model ignored
  - With life loss model, BESS operation optimized to minimize cost and ensure expected lifetime
  - System with the life loss model considered offers 10% more savings over the calendar life



Model	$\bar{l}_t^{cyc}$ and $\bar{l}_t^{cal}$	Expected lifetime (years)
Without life loss	-	1.7
With life loss	1.1e-3	2.5
With life loss	9.1e-4	3

Conclusion and Future Work

- This work proposes an innovative DDPG-based RL approach for optimal DER dispatch, considering BESS calendrical and cyclical life loss models.
- Case studies were performed to show that the proposed approach can maximize benefits while ensuring the BESS lifespan requirement.
- One interesting future work is to develop an advanced RL approach based on the policy iteration strategy for DER coordination considering distribution power flow.
- Another interesting direction is to develop advanced multi-objective optimization to find a good balance between economic benefits and battery lifetime.

* A. Das and D. Wu. "Optimal Coordination of Distributed Energy Resources Using Deep Deterministic Policy Gradient." in *Proc. IEEE Power Energy Soc. Electr. Energy Stor. App. and Tech. Conf. (EESAT)*, Austin, TX, Nov. 2022, pp. 1-5.

* Y. Du and D. Wu, "Deep reinforcement learning from demonstrations to assist service restoration in islanded microgrids," *IEEE Trans. Sustain. Energy*, vol. 13, no. 2, pp. 1062–1072, Apr. 2022.

* A. Das, D. Wu, and B.A. Bhatti, "Approximate Dynamic Programming with Enhanced Off-policy Learning for Coordinating Distributed Energy Resources," *IEEE Trans. Sustain. Energy*, Jul. 2023. (To Be Submitted)

Acknowledgment

This work is supported by the U.S. Department of Energy, Office of Electricity through Energy Storage Program. We are grateful to **Dr. Imre Gyuk** for providing financial support and leadership on this and other related work at Pacific Northwest National Laboratory.

Thank You

Avijit Das
avijit.das@pnnl.gov

Backup Slide: DDPG Technique

- Critic network
  - Estimate the action-value function, and the network weights are updated to minimize the mean square error

Estimate

$$E_c = \frac{1}{M}\sum_{m=1}^{M}(e_c^m)^2$$

Target

$$e_c^m = Q\left(S_t^m, x_t^m \middle| \theta_Q\right) - \left(r_t^m + \gamma Q'\left(S_{t+1}^m, x_{t+1}^m \middle| \theta_Q'\right)\right)$$

$$\theta_Q = \theta_Q - \Phi_c \nabla_{\theta_Q} E_c;\ \theta_{Q'} = \tau\theta_Q - (1-\tau)\theta_{Q'}$$

Soft update

- Actor network
  - Learn a policy that minimizes the action-value function

$$\theta_v = \theta_v - \Phi_a \nabla_{\theta_v} J(\theta_v);\ \theta_{v'} = \tau\theta_v - (1-\tau)\theta_{v'}$$

$$\nabla_{\theta_v} J(\theta_v) = \frac{1}{M}\sum_{m=1}^{M}\nabla_{x_t^m} Q\left(S_t^m, x_t^m \middle| \theta_Q\right) \nabla_{\theta_v} v\left(S_t^m \middle| \theta_v\right)$$

- Safe exploration is ensured through action mapping, clipping, and reward shaping