

Control-Guided Reinforcement Learning for Cooperative Energy Management

Isabela Fons^{*1}, Raquel Salcedo¹, Rubén Ruiz¹, José A. Caballero¹, Antonio del Río Chanona²
 University of Alicante¹ | Imperial College London²
 * isabela.fons@ua.es



MOTIVATION

Microgrid → small-scale energy production and distribution network that can operate independently or in tandem with the traditional grid. Key features:

- decentralized and stochastic
- key in path to decarbonization

Classical control → limited scalability

RL-based control → promising but sample inefficient

RESEARCH GAP: Existing MARL approaches do not integrate expert knowledge, despite its success in other control domains [1].



Cost minimization via temperature and battery control under comfort and mobility constraints.

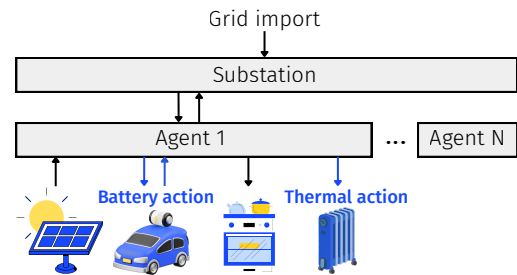


Figure 1 – Schematic representation of an energy microgrid [2, 3]

CASE STUDY

The proposed approach combines BC and CEM for decentralized coordination.

- **Behavior Cloning (BC)** → policy warm-start with imitation learning from PID controller
- **Cross-Entropy Method (CEM)** → policy fine tuning using derivative-free optimization
- **Independent learning** → multiple agents operate simultaneously within a shared environment, but policies are updated individually based on local information

METHODOLOGY

Training behavior

- CEM successfully learns control policies under uncertainty (weather conditions and domestic loads)
- BC initialization reaches a higher performance faster

Figure 3 – Learning curves for CEM and BC-CEM

Operational insights

RL policies preserve smoothness of PID control and exploit system flexibility.

Figure 4 – Temperature evolution of an example day

RESULTS

CONCLUSION

Combining control knowledge and derivative-free MARL is a powerful control strategy for complex energy systems

- **Improved** learning speed and policy **performance**
- **Reduced exploration burden** by incorporating expert experience
- Pathway to **scalable decentralized control**

REFERENCES

- [1] Bloor et al., Ind. Eng. Chem. Res. (2025)
- [2] Charbonnier et al., Appl. Energy (2023)
- [3] Charbonnier et al., Appl. Energy (2025)

ACKNOWLEDGEMENTS

Supported by Spanish Ministry of Science and Innovation (fellowship FPU22/01276, projects PID2021-124139NB-C21 and PID2024-155232NB-I00).

Follow this research and future updates



GitHub



LinkedIn

Evaluation (200 test episodes)

- Up to 180% cost reduction vs PID
- Comparable battery availability
- Comfort deviations remain limited and within acceptable bounds

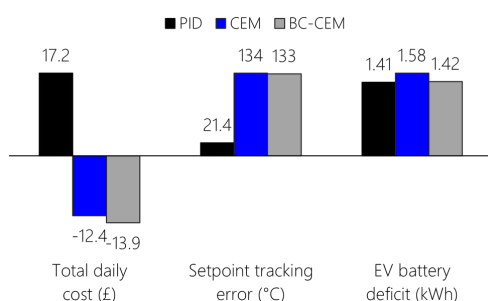


Fig. 5 – Mean daily results for a 10-prosumer microgrid

- Outperforms all baselines in average return

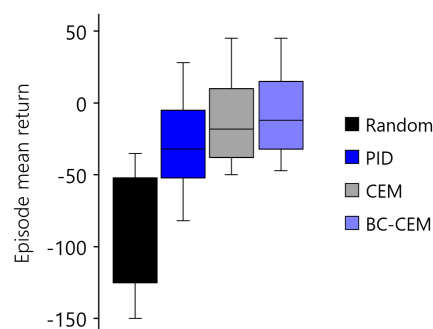


Fig. 6 – Evaluation episode return comparison