Tackling Environment Sustainability Challenges via Reinforcement Learning and Counterfactual Explanations

Luca Marzari[1] **and Enrico Marchesini**[2]

[1]University of Verona, Verona, Italy
[2]Massachusetts Institute of Technology, Cambridge (MA), USA
luca.marzari@univr.it, emarche@mit.edu

Abstract

Reinforcement learning (RL) offers transformative potential for many environmental sustainability problems ranging from water monitoring with data acquisition to power grid operations, by learning adaptive and scalable controllers through trial and error. However, current RL methods face significant challenges when applied to real-world settings, including handling the uncertainty of complex system dynamics, achieving long-term objectives, and satisfying strict physical constraints. Moreover, the lack of transparency of RL decisions, arising from the black-box nature of deep neural networks, is hindering the deployment of RL in our society. Progress in this field requires close collaboration with environmental operators to ensure RL methods are effective, explainable, and especially safe. To this end, we argue that proposing innovative combinations of explainable AI solutions, such as counterfactual explanations, with recent RL approaches can be an interesting direction to consider and pave the way to trustworthy and effective control solutions for environmental operations.

Motivations. Many environmental sustainability problems play an important role in combating climate change, necessitating a rapid transition to sustainable practices that enhance resilience to climate-driven extreme events (European Commission 2024). This transition places new demands on the systems that should now operate more dynamically, on larger scales, and under increased uncertainty. Much of this complexity stems from shifting resource availability and usage patterns, driven by integrating renewable energy sources, distributed monitoring technologies, and evolving environmental factors (Li et al. 2023). These changes introduce significant challenges for human operators and reveal the limitations of traditional analytical and management tools in effectively addressing the complexities of modern systems (Marot et al. 2022). Reinforcement learning (RL) holds considerable promise for revolutionizing problems ranging from water monitoring, data acquisition, and power grid operations, as evidenced by its success in complex game scenarios (Mnih et al. 2013; Papoudakis et al. 2021). However, several challenges remain in applying RL to real-world tasks, including the need to manage system dynamics, account for aleatoric uncertainty, pursue long-term objectives, and adhere to stringent physical constraints. Among other problems, power grids embody many of these challenges, which are also active research areas within the RL field. Thus, *exploring realistic power grid tasks through an RL lens could drive valuable progress in both RL research and societal needs*. Deploying RL-based solutions to these complex physical systems, however, requires collaborating with system operators and providing them with trustworthy and explainable solutions. This aspect is seldom straightforward to achieve due to the black-box nature of deep neural networks on which scalable RL solutions are built upon.

Explainable RL solutions for power grid operations. Our solutions to this open challenge comprise the ongoing development of *RL2Grid*, an RL framework that models realistic power grid operations and captures a range of increasingly complex tasks that received an honorable mention at the MIT Prize for Open Data (MIT Libraries 2024). This framework mirrors the combinatorially large action space typical of grid operations and is presented within a standardized Gymnasium-based interface, complete with shared reward structures, state spaces, and action spaces to enable standardized evaluation for future developments. While deep RL methods hold promise for addressing many pressing challenges in power grid management, they also introduce potential risks—such as those related to safety, reliability, and robustness—that require careful consideration. Future research directions in deep RL for power grid applications include: (i) safe RL methods to ensure learning and control policies prevent actions that might lead to blackouts or equipment damage; (ii) human-in-the-loop systems, which integrate human supervision, interaction, and feedback for a synergistic approach where human operators and AI collaborate to optimize grid operations; (iii) novel and scalable verification/explainable deep neural network methods to enhance the understanding of the RL decisions (Jiang et al. 2024; Marzari et al. 2024). Specifically, counterfactual explanations for RL (Gajcin and Dusparic 2024) hold great promise in providing an explainable decision-making process for RL policies. However, given the NP-hardness of formal verification of neural networks, the necessity of discovering novel scalable and efficient solutions for real-world environmental sustainability challenges remains an attractive open problem we aim to address in the next years.

References

European Commission. 2024. Cluster 6: Food, Bioeconomy, Natural Resources, Agriculture and Environment. https://research-and-innovation.ec.europa.eu/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-europe/cluster-6-food-bioeconomy-natural-resources-agriculture-and-environment_en.

Gajcin, J.; and Dusparic, I. 2024. Redefining Counterfactual Explanations for Reinforcement Learning: Overview, Challenges and Opportunities. *ACM Computing Surveys*, 56(9): 1–33.

Jiang, J.; Leofante, F.; Rago, A.; and Toni, F. 2024. Robust counterfactual explanations in machine learning: A survey. *IJCAI*.

Li, Y.; Yu, C.; Shahidehpour, M.; Yang, T.; Zeng, Z.; and Chai, T. 2023. Deep Reinforcement Learning for Smart Grid Operations: Algorithms, Applications, and Prospects. *Proceedings of the IEEE*, 111(9): 1055–1096.

Marot, A.; Kelly, A.; Naglic, M.; Barbesant, V.; Cremer, J.; Stefanov, A.; and Viebahn, J. 2022. Perspectives on future power system control centers for energy transition. *Journal of Modern Power Systems and Clean Energy*, 10(2): 328–344.

Marzari, L.; Leofante, F.; Cicalese, F.; and Farinelli, A. 2024. Rigorous Probabilistic Guarantees for Robust Counterfactual Explanations. In *ECAI 2024*, 1059–1066. IOS Press.

MIT Libraries. 2024. MIT Prize for Open Data. https://libraries.mit.edu/opendata/open-data-mit-home/mit-prize/.

Mnih, V.; Kavukcuoglu, K.; Silver, D.; Graves, A.; Antonoglou, I.; Wierstra, D.; and Riedmiller, M. 2013. Playing Atari with Deep Reinforcement Learning. In *Conference on Neural Information Processing Systems (NeurIPS)*.

Papoudakis, G.; Christianos, F.; Schäfer, L.; and Albrecht, S. V. 2021. Comparative Evaluation of Multi-Agent Deep Reinforcement Learning Algorithms. In *Conference on Neural Information Processing Systems Datasets and Benchmarks Track (NeurIPS)*.