

# Multi-Objective Reinforcement Learning for Water Management

Extended Abstract

Zuzanna Osika  
Delft University of Technology  
Delft, Netherlands  
z.osika@tudelft.nl

Roxana Rădulescu  
Utrecht University  
Utrecht, Netherlands  
r.t.radulescu@uu.nl

Jazmin Zatarain-Salazar  
Delft University of Technology  
Delft, Netherlands  
j.zatarain-salazar@tudelft.nl

Frans A. Oliehoek  
Delft University of Technology  
Delft, Netherlands  
f.a.oliehoek@tudelft.nl

Pradeep K. Murukannaiah  
Delft University of Technology  
Delft, Netherlands  
p.k.murukannaiah@tudelft.nl

## ABSTRACT

Many real-world problems (e.g., resource management, autonomous driving, drug discovery) require optimizing multiple, conflicting objectives. Multi-objective reinforcement learning (MORL) extends classic reinforcement learning to handle multiple objectives simultaneously, yielding a set of policies that capture various trade-offs. However, the MORL field lacks complex, realistic environments and benchmarks. We introduce a water resource (Nile river basin) management case study and model it as a MORL environment. We then benchmark existing MORL algorithms on this task. Our results show that specialized water management methods outperform state-of-the-art MORL approaches, underscoring the scalability challenges MORL algorithms face in real-world scenarios.

## CCS CONCEPTS

• **Theory of computation** → **Sequential decision making**.

## KEYWORDS

Multi-Objective Reinforcement Learning, Water Management

### ACM Reference Format:

Zuzanna Osika, Roxana Rădulescu, Jazmin Zatarain-Salazar, Frans A. Oliehoek, and Pradeep K. Murukannaiah. 2025. Multi-Objective Reinforcement Learning for Water Management: Extended Abstract. In *Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Detroit, Michigan, USA, May 19 – 23, 2025*, IFAAMAS, 3 pages.

## 1 INTRODUCTION

Multi-objective reinforcement learning (MORL) extends the classic reinforcement learning (RL) framework [20] to problems with multiple, conflicting objectives. Rather than maximizing a single reward, MORL uses a vector of rewards—one per objective—yielding multiple optimal policies with distinct trade-offs [6]. It has been applied in water management [3, 16], autonomous driving [9], power allocation [12, 23], drone navigation [22], and medical treatment

[8, 10]. Recent advances [1, 2, 24] have expanded MORL’s algorithmic capabilities, yet benchmarking still heavily relies on toy tasks. As a result, the real-world applicability of state-of-the-art algorithms remains unclear, especially given the additional objective dimension alongside large state and action spaces.

Water resource management is becoming increasingly complex due to climate change. A key problem in this domain is about making sequential decisions about water releases from dams to balance multiple, often conflicting objectives, including hydropower generation, irrigation, water supply, and environmental preservation. This multi-objective problem has been studied in real-world contexts [4, 15, 16, 25] using methods such as dynamic programming and mathematical optimization [5, 7]. A specialized method, evolutionary multi-objective direct policy search (EMODPS) [3], has demonstrated strong performance [17, 19]. However, existing models of the problem remain difficult to access due to proprietary data, multiple programming languages, and inconsistent code structures.

In this paper, we address the problem of the lack of large-scale, real-world problems for evaluating MORL algorithms. We introduce a real-world case study from the field of water resource management and model it in a modular, structured way as a MORL environment using MO-Gymnasium [2] API. This case study enables MORL researchers to test their algorithms in realistic, high-impact scenarios. The case study we introduce is Nile river basin management. By benchmarking state-of-the-art MORL algorithms in this environment, we evaluate their performance against the specialized water management algorithm, EMODPS, and explore the effectiveness of existing MORL methods for real-world application.<sup>1</sup>

## 2 BACKGROUND

We formalize a MORL problem as a multi-objective Markov decision process (MOMDP)  $\langle S, A, T, \gamma, \mu, \mathbf{R} \rangle$ , where  $S$  and  $A$  are the state and action spaces,  $T$  is a probabilistic transition function,  $\gamma$  is a discount factor,  $\mu$  is the initial state distribution, and  $\mathbf{R}$  is a vector-valued reward function returning a numeric feedback signal for each of  $d \geq 2$  objectives—the key difference from standard MDPs.

Because the policy value  $\mathbf{V}^\pi \in \mathbb{R}^d$  can only be partially ordered, either a utility function  $u : \mathbb{R}^d \rightarrow \mathbb{R}$  is used to prioritize objectives or, if unknown, an entire solution set (multi-policy paradigm) must be returned. The **Pareto set** is the set of all pairwise undominated policies based on Pareto dominance, while the **Pareto front** (PF) is

<sup>1</sup>Code: [https://github.com/osikazuzanna/wms\\_morl/](https://github.com/osikazuzanna/wms_morl/)



This work is licensed under a Creative Commons Attribution International 4.0 License.

the corresponding set of value vectors, forming the **convex hull** (CH) if  $u$  is a positively weighted linear sum [13].

### 3 NILE RIVER BASIN

**Background** The Nile River, a vital resource for ten northeast African countries, supports hydropower, agriculture, and municipal needs but faces ongoing political tensions over water rights, particularly between Egypt, Sudan, and Ethiopia. Ethiopia’s Grand Ethiopian Renaissance Dam (GERD), seen as essential for economic development and expanding electricity access (currently only 27%), is a major point of contention. Egypt, historically reliant on the Nile and the High Aswan Dam (HAD) for hydropower and water security, fears GERD will reduce downstream flows, jeopardizing food and energy security. Sudan, benefiting from extensive irrigation infrastructure, initially opposed GERD alongside Egypt but now acknowledges potential flood regulation benefits while still expressing concerns about water security. Negotiations remain stalled as the river faces additional challenges, including variable inflows, droughts, floods, and increasing demand due to population growth [21]. Our study models the Nile River Basin to capture the conflicting interests of Ethiopia, Sudan, and Egypt.

**Actions** The action tuple is four-dimensional, where each dimension represents the percentage of water to be released from one of the following dams: GERD, Roseires, Sennar, and HAD. The actions are taken every month, for 20 years. Thus, the full episode consists of 240 timesteps.

**Observations** The state is five-dimensional, with the components representing the storage of the four reservoirs modeled and the month of the year the system is in (normalized).

**Rewards** We consider four objectives; so, the reward is four-dimensional: Egypt irrigation deficit (ED), Sudan irrigation deficit (SD), Egypt minimum HAD level, which represents water reliability in the HAD ensuring that the water level exceeds the minimum required for power generation (HAD), and Ethiopia hydropower generation from the GERD power plant (EH).

### 4 EXPERIMENTS AND RESULTS

**Experimental Setup** We benchmark three state-of-the-art multi-policy MORL algorithms, covering both CH and PF-returning methods: GPI-LS [1], PCN [18] and CAPQL [11]. These algorithms were chosen specifically as they can handle continuous actions and state space as well as more than two objectives. These are compared with EMODPS [3], an advanced optimization approach for designing efficient control policies in complex water management systems. To compare the performance of these algorithms we use hypervolume (region between the solution set and a reference point, which defines the lower bound for each objective) ( $\uparrow$ ), cardinality (number of solutions in the final set) ( $\nearrow$ ) and sparsity (spread of the solutions in the solution set) ( $\downarrow$ ) of the solution set [24]. We trained each MORL algorithm for 200,000 timesteps and EMODPS for 20,000 NFEs, selecting these values based on observed convergence.

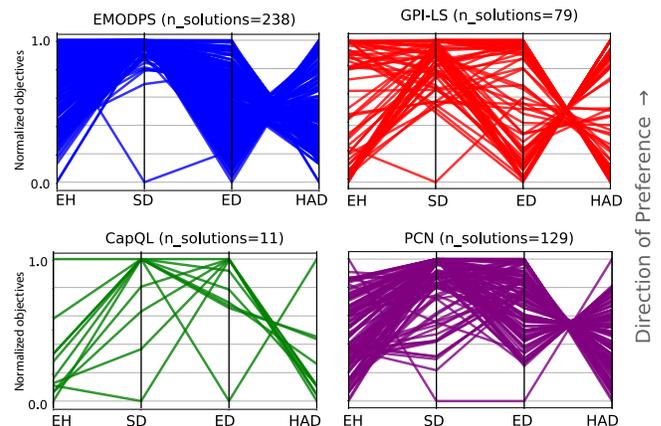
**Results** Table 1 shows final mean metrics (over five seeds) and uses EMODPS hypervolume as the baseline (value=1 indicates matching performance). GPI-LS achieves the highest MORL hypervolume (68% of the baseline), while PCN and CapQL remain at 10–11%.

EMODPS outperforms in cardinality, which measures the number of solutions in the final set. Despite fewer solutions and higher sparsity, GPI-LS covers diverse trade-offs, whereas PCN and CapQL concentrate on narrower solution regions.

Algorithm	Hypervolume (% Baseline) $\uparrow$	Cardinality $\nearrow$	Sparsity $\downarrow$
EMODPS (Baseline)	2.21E+08 (100%)	327	3.951
PCN	2.26E+07 (10%)	51	7.754
GPI-LS	1.50E+08 (68%)	41	150.636
CapQL	2.03E+06 (1%)	9	7.719

**Table 1: Comparison of MORL and EMODPS algorithms**

Figure 1 shows parallel coordinate plots [14] of each algorithm’s solutions, with the x-axis listing objectives and the y-axis indicating normalized objective values (higher is better); each polyline represents a solution, and an ideal one would be a flat line at 1.0.



**Figure 1: Parallel coordinate plots with solution sets achieved by each algorithm for the Nile environment.**

The solution sets for each algorithm were generated by merging results from multiple seeds and applying Pareto filtering to retain only non-dominated solutions, with counts shown above each plot. PCN and GPI-LS produce larger merged sets than their average cardinality, suggesting distinct regional coverage across seeds, making exploration an important challenge in this domain. EMODPS achieves a similar size to its mean cardinality, indicating comprehensive exploration within each seed.

### 5 CONCLUSIONS

We emphasize the importance of applying MORL algorithms to real-world challenges, as current approaches, often tested on theoretical problems, struggle with practical issues like exploration and scalability. By introducing a complex water management case, we hope to encourage the MORL field to focus on practical challenges when developing new algorithms. Our results show that current MORL algorithms fall short in comparison to the domain-specific water resource management algorithm, EMODPS. We also highlight two challenges for MORL approaches: scalability and exploration.

## REFERENCES

- [1] Lucas N. Alegre, Ana L. C. Bazzan, Diederik M. Roijers, Ann Nowé, and Bruno C. da Silva. 2023. Sample-Efficient Multi-Objective Learning via Generalized Policy Improvement Prioritization. In *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems* (London, United Kingdom) (AMAS '23). International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 2003–2012.
- [2] Florian Felten, Lucas N. Alegre, Ann Nowe, Ana Bazzan, El Ghazali Talbi, Grégoire Danoy, and Bruno C. da Silva. 2023. A Toolkit for Reliable Benchmarking and Research in Multi-Objective Reinforcement Learning. In *Advances in Neural Information Processing Systems*, A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine (Eds.), Vol. 36. Curran Associates, Inc., 23671–23700. [https://proceedings.neurips.cc/paper\\_files/paper/2023/file/4aa8891583f07ae200ba07843954caeb-Paper-Datasets\\_and\\_Benchmarks.pdf](https://proceedings.neurips.cc/paper_files/paper/2023/file/4aa8891583f07ae200ba07843954caeb-Paper-Datasets_and_Benchmarks.pdf)
- [3] Matteo Giuliani, Andrea Castelletti, Francesca Pianosi, Emanuele Mason, and Patrick M Reed. 2016. Curses, tradeoffs, and scalable management: Advancing evolutionary multiobjective direct policy search to improve water reservoir operations. *Journal of Water Resources Planning and Management* 142, 2 (2016), 04015050.
- [4] Matteo Giuliani, Jonathan D Herman, Andrea Castelletti, and P Reed. 2014. Many-objective reservoir policy identification and refinement to reduce policy inertia and myopia in water management. *Water resources research* 50, 4 (2014), 3355–3377.
- [5] M Giuliani, JR Lamontagne, PM Reed, and A Castelletti. 2021. A state-of-the-art review of optimal reservoir control for managing conflicting demands in a changing world. *Water Resources Research* 57, 12 (2021), e2021WR029927.
- [6] Conor F Hayes, Roxana Rădulescu, Eugenio Bargiacchi, Johan Källström, Matthew Macfarlane, Mathieu Reymond, Timothy Verstraeten, Luisa M Zintgraf, Richard Dazeley, Fredrik Heintz, et al. 2022. A practical guide to multi-objective reinforcement learning and planning. *Autonomous Agents and Multi-Agent Systems* 36, 1 (2022), 26.
- [7] Jonathan D Herman, Julianne D Quinn, Scott Steinschneider, Matteo Giuliani, and Sarah Fletcher. 2020. Climate adaptation as a control problem: Review and perspectives on dynamic water resources planning under uncertainty. *Water Resources Research* 56, 2 (2020), e24389.
- [8] Ammar Jalalimanesh, Hamidreza Shahabi Haghighi, Abbas Ahmadi, Hossein Hejazian, and Madjid Soltani. 2017. Multi-objective optimization of radiotherapy: distributed Q-learning and agent-based simulation. *Journal of Experimental & Theoretical artificial intelligence* 29, 5 (2017), 1071–1086.
- [9] Changjian Li and Krzysztof Czarnecki. 2019. Urban Driving with Multi-Objective Deep Reinforcement Learning. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems* (Montreal QC, Canada) (AMAS '19). International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 359–367.
- [10] Daniel J. Lizotte, Michael Bowling, and Susan A. Murphy. 2010. Efficient reinforcement learning with multiple reward functions for randomized controlled trial analysis. In *Proceedings of the 27th International Conference on International Conference on Machine Learning* (Haifa, Israel) (ICML '10). Omnipress, Madison, WI, USA, 695–702.
- [11] H. Lu, D. Herman, and Y. Yu. 2023. Multi-Objective Reinforcement Learning: Convexity, Stationarity and Pareto Optimality. In *International Conference on Learning Representations (ICLR)*. Curran Associates, Inc. <https://openreview.net/forum?id=TjEzlsyEsQ6>
- [12] Youngwoo Oh, Arif Ullah, and Wooyeol Choi. 2024. Multi-Objective Reinforcement Learning for Power Allocation in Massive MIMO Networks: A Solution to Spectral and Energy Trade-Offs. *IEEE Access* 12 (2024), 1172–1188. <https://doi.org/10.1109/ACCESS.2023.3347788>
- [13] Zuzanna Osika, Jazmin Zatarain-Salazar, Frans A Oliehoek, and Pradeep K Murukannaiah. 2024. Navigating Trade-offs: Policy Summarization for Multi-Objective Reinforcement Learning. In *ECAI 2024*. IOS Press, 2919–2926.
- [14] Zuzanna Osika, Jazmin Zatarain Salazar, Diederik M. Roijers, Frans A. Oliehoek, and Pradeep K. Murukannaiah. 2023. What Lies beyond the Pareto Front? A Survey on Decision-Support Methods for Multi-Objective Optimization. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, IJCAI-23*, Edith Elkind (Ed.). International Joint Conferences on Artificial Intelligence Organization, 6741–6749. <https://doi.org/10.24963/ijcai.2023/755>
- [15] Afua Owusu, Jazmin Zatarain Salazar, Marloes Mul, Pieter van der Zaag, and Jill Slinger. 2022. Quantifying the trade-offs in re-operating dams for the environment in the Lower Volta River. *Hydrology and Earth System Sciences Discussions* 2022 (2022), 1–27.
- [16] F. Pianosi, A. Castelletti, and M. Restelli. 2013. Tree-based fitted Q-iteration for multi-objective Markov decision processes in water resource management. *Journal of Hydroinformatics* 15, 2 (01 2013), 258–270. <https://doi.org/10.2166/hydro.2013.169> arXiv:<https://iwaponline.com/jh/article-pdf/15/2/258/386917/258.pdf>
- [17] Julianne D Quinn, Patrick M Reed, Matteo Giuliani, and Andrea Castelletti. 2019. What is controlling our control rules? Opening the black box of multi-reservoir operating policies using time-varying sensitivity analysis. *Water Resources Research* 55, 7 (2019), 5962–5984.
- [18] Mathieu Reymond, Eugenio Bargiacchi, and Ann Nowe. 2022. Pareto Conditioned Networks. In *The 21st International Conference on Autonomous Agents and Multiagent Systems*. IFAAMAS, 1110–1118. <https://aamas2022-conference.auckland.ac.nz>
- [19] Jazmin Zatarain Salazar, Patrick M Reed, Jonathan D Herman, Matteo Giuliani, and Andrea Castelletti. 2016. A diagnostic assessment of evolutionary algorithms for multi-objective surface water reservoir control. *Advances in water resources* 92 (2016), 172–185.
- [20] Richard S. Sutton and Andrew G. Barto. 2018. *Reinforcement Learning: An Introduction* (second ed.). The MIT Press. <http://incompleteideas.net/book/the-book-2nd.html>
- [21] Kevin G Wheeler, Jim W Hall, Gamal M Abdo, Simon J Dadson, Joseph R Kasprzyk, Rebecca Smith, and Edith A Zagona. 2018. Exploring cooperative transboundary river management strategies for the Eastern Nile Basin. *Water Resources Research* 54, 11 (2018), 9224–9254.
- [22] Jiahao Wu, Yang Ye, and Jing Du. 2024. Multi-objective reinforcement learning for autonomous drone navigation in urban areas with wind zones. *Automation in Construction* 158 (2024), 105253.
- [23] Zheng Xiong, Biao Luo, Bing-Chuan Wang, Xiaodong Xu, and Tingwen Huang. 2024. Multiobjective Battery Charging Strategy Based on Deep Reinforcement Learning. *IEEE Transactions on Transportation Electrification* 10, 3 (2024), 6893–6903. <https://doi.org/10.1109/TTE.2023.3347297>
- [24] Jie Xu, Yunsheng Tian, Pingchuan Ma, Daniela Rus, Shinjiro Sueda, and Wojciech Matusik. 2020. Prediction-guided multi-objective reinforcement learning for continuous robot control. In *Proceedings of the 37th International Conference on Machine Learning (ICML '20)*. JMLR.org, Article 983, 10 pages.
- [25] Marta Zaniolo, Matteo Giuliani, and Andrea Castelletti. 2021. Neuro-evolutionary direct policy search for multiobjective optimal control. *IEEE Transactions on Neural Networks and Learning Systems* 33, 10 (2021), 5926–5938.