

Staged Design of Water Distribution Networks: A Reinforcement Learning Approach [†]

Lydia Tsiami ^{1,2,*} , Christos Makropoulos ^{1,2} and Dragan Savic ^{1,3,4} 

¹ KWR Water Research Institute, 3433 PE Nieuwegein, The Netherlands; cmakro@mail.ntua.gr (C.M.); dragan.savic@kwrwater.nl (D.S.)

² Department of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical University of Athens, 157 80 Athens, Greece

³ Centre for Water Systems, University of Exeter, Exeter EX4 4PY, UK

⁴ Faculty of Civil Engineering, University of Belgrade, 11000 Belgrade, Serbia

* Correspondence: lydia.tsiami@kwrwater.nl

[†] Presented at the 3rd International Joint Conference on Water Distribution Systems Analysis & Computing and Control for the Water Industry (WDSA/CCWI 2024), Ferrara, Italy, 1–4 July 2024.

Abstract: Effectively planning the design of a water distribution network for the long term is a challenging task for water utilities, mainly due to the deep uncertainty that characterizes some of its most important design parameters. In an effort to navigate this challenge, this work investigates the potential of reinforcement learning in the lifecycle design of water networks. To this end, a deep reinforcement learning agent was trained to identify a sequence of cost-effective interventions across multiple construction phases within a network's lifecycle under both deterministic and uncertain conditions. Our approach was tested on a modified benchmark of the New York Tunnels problem with promising results. The agent achieved comparable performance with the baseline heuristic algorithm in the deterministic setting and devised a flexible design strategy when multiple future scenarios were considered. These preliminary findings highlight the potential of reinforcement learning in the lifecycle design of water networks and represent a step towards the integration of more adaptive planning approaches in the field.

Keywords: flexibility; robustness; design; long-term planning; reinforcement learning



Citation: Tsiami, L.; Makropoulos, C.; Savic, D. Staged Design of Water Distribution Networks: A Reinforcement Learning Approach. *Eng. Proc.* **2024**, *69*, 111. <https://doi.org/10.3390/engproc2024069111>

Academic Editors: Stefano Alvisi, Marco Franchini, Valentina Marsili and Filippo Mazzoni

Published: 10 September 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Effectively planning the design of water distribution networks (WDNs) for the long term is an essential but challenging task for water utilities. Construction interventions need to be carefully planned within budget constraints and designed to meet both present and future needs, despite uncertainties in variables such as population growth and climate change. These uncertain variables, which are essential for describing the WDN in the future, are characterized by “deep uncertainty” [1]. This introduces an additional layer of complexity to the problem. To better navigate these challenges, it is essential to move towards adaptive planning techniques that adjust to emerging information.

Reinforcement learning (RL), a subfield of machine learning focused on sequential decision-making within uncertain and dynamic environments, could be a promising solution to the problem. Unlike heuristic approaches, RL offers a more dynamic framework for decision-making under deep uncertainty as it can adapt to emerging information [2].

This work builds upon our prior research demonstrating RL's feasibility in the single-stage WDN design [3]. Here, we explore the more challenging staged optimization problem [4]. To this end, a deep RL agent was trained to identify a sequence of cost-effective interventions across multiple construction phases within a network's lifecycle. Our approach was tested on a modified benchmark of the New York Tunnels (NYT) problem, under deterministic and uncertain conditions. By comparing the performance of our agent with heuristic algorithms, we discuss the potential of RL in the lifecycle design of WDNs.

2. Preliminaries

Staged Design Approaches

Staged design is a dynamic approach for the lifecycle design of WDN, in which the planning horizon is divided into multiple construction stages, each one with its own requirements. Within each stage, a set of specific upgrades are identified and implemented to meet immediate demands while also optimizing an overall objective—typically the total present cost of the network—across the entire planning horizon. This phased approach allows water utilities to prioritize current needs without neglecting the network's future growth [4]. Additionally, by deferring some costs to later stages, utilities can manage investments more effectively, avoiding both high initial expenditures and costly, reactive interventions needed to cope with unforeseen future demands.

In the literature, there are three staged design variations [4]: (1) deterministic approaches, which assume that demands are known for the whole planning horizon; (2) robust approaches, aiming to identify a solution that works well enough for a range of scenarios; and (3) flexible approaches, focusing on identifying a set of initial interventions that will allow the network to adapt to different scenarios with few modifications.

3. Materials and Methods

Using RL for Lifecycle Design of Water Distribution Networks

We explore the feasibility of RL in the lifecycle design of WDNs, focusing on the single-objective staged optimization problem. The problem involves optimizing pipe diameters across multiple construction stages ($k \in K$) in a network of n nodes and m pipes to minimize the total cost while meeting stage-specific demands and pressure requirements.

We start by formulating the problem as a Markov Decision Process (MDP) defined by the tuple (S, A, T, R) . The problem is mapped to an MDP as follows:

- State: A vector s_t containing the pressure at each node and the diameters of the pipes.
- Action: A tuple $a_t = (a_{1t}, a_{2t})$ where $a_{1t} \in A_1$, a discrete set of all network pipes, and $a_{2t} \in A_2$, a discrete set of all commercially available pipe diameters.
- Transition function: T is deterministic when analyzing a single scenario and stochastic when designing for multiple future scenarios simultaneously.
- Reward function: A combination of two terms: (1) the cost of the design and (2) a penalty term proportional to the total pressure deficit in the network.

We express the staged cost-minimization problem as a sequential decision-making task. An agent, starting with an initial network design, takes $L = \sum_{k=1}^K l_k$ number of actions, where l_k is the number of upgrades at construction stage k , to determine an optimal sequence of interventions that minimize the total cost while pressure requirements are met. For this task, we use the Proximal Policy Optimization (PPO) algorithm [5].

4. Case Study

We apply our methodology in an adapted case study of NYT. The original problem involved reinforcing the network to meet demands with enough pressure and under a single demand condition [6]. In the adapted version of this problem, Cunha et al. [7] transformed the problem into a staged optimization problem considering a 60-year lifespan and three construction stages (every 20 years). At each stage, there is a uniform demand increase (0–10%) and the Hazen–Williams coefficient decreases by 2.5 every decade. The goal is to minimize the total intervention cost assuming a 4% discount rate.

Cunha et al. generated 50 possible scenarios and their objective was to find a robust staged solution that worked well enough for all of them. To achieve that, they first identified the cost-optimal solutions for 10 reference scenarios. Then, they evaluated and ranked each solution under all generated scenarios using multicriteria decision analysis.

For our experiment, we focused on the 10 reference scenarios and their corresponding deterministic solutions to assess our agent's ability to solve staged optimization problems.

We then further challenged the agent by testing its ability to develop a flexible strategy that can adapt to all 10 scenarios. The results are presented in the next section.

5. Results

5.1. Staged Optimization

Table 1 compares the RL agent's performance with the heuristic algorithm across the 10 benchmark scenarios. The agent achieved comparable total cost performance throughout the network's lifecycle, successfully identifying feasible and cost-effective solutions. The agent's designs were slightly more expensive compared to the designs generated by the heuristic algorithm (0.6% to 5.1%). However, it is also worth noting that for scenario 10, the agent was able to find a marginally more cost-effective solution.

Table 1. Cost comparison of staged solutions. Baseline heuristic algorithm [7] vs. PPO agent.

Scenario	Demand Increase (%)			Total Demand Increase (%)	Cost (Heuristic [7]) ($\times 10^7$ USD)	Cost (PPO Agent) ($\times 10^7$ USD)
	t = 0	t = 20	t = 40			
1	0	0	0	0	4.09	4.17
2	0	3	4	7.12	4.54	4.64
3	0	5	1	6.05	4.63	4.71
4	5	1	1	7.11	5.14	5.40
5	4	3	3	10.33	5.26	5.29
6	3	3	9	15.64	5.34	5.51
7	9	1	3	13.39	6.07	6.31
8	8	6	1	15.62	6.25	6.44
9	6	9	9	25.94	6.42	6.46
10	10	10	10	33.10	7.55	7.48

5.2. Flexible Optimization

Finding a flexible design that can adapt to several scenarios with few modifications requires a common starting point. Given that the 10 reference scenarios had different initial demands, we added a common initial stage for all scenarios with demand equal to the demand of the original NYT problem (Table 2). During training, the agent encountered each scenario an equal number of times but in a random order. As a result, the agent developed a flexible and cost-efficient strategy that applies to all scenarios.

Table 2. Cost comparison of scenario-optimal solutions against the flexible ones.

Scenario	Demand Increase (%)				Deterministic Cost ($\times 10^7$ USD)	Flexible Cost ($\times 10^7$ USD)
	t = 0	t = 20	t = 40	t = 60		
1	0	0	0	0	4.33	4.92
2	0	0	3	4	4.57	4.97
3	0	0	5	1	4.51	5.01
4	0	5	1	1	5.00	5.15
5	0	4	3	3	5.07	5.12
6	0	3	3	9	4.87	5.08
7	0	9	1	3	5.44	5.73
8	0	8	6	1	5.34	5.78
9	0	6	9	9	5.36	5.51
10	0	10	10	10	5.93	6.08
Avg. Cost					-	5.33

The agent began with a common network design for all ten scenarios, and then progressively adapted the network to meet the demand requirements of each scenario. Interestingly, in the early construction stages, when the uncertainty about future demand was higher, the agent divided the scenarios into subgroups and applied the same actions for each subgroup. Then, as the planning process progressed and the scenarios began to diverge more significantly, the agent's strategy became more tailored to each scenario. For instance, instead of implementing one unique intervention for each scenario at each stage, the second stage had just two possible interventions; one for scenarios with demand

increases below a certain threshold and another for those above it. The third stage had 7 unique actions, and then finally, the fourth stage had 10 (1 for each scenario).

Table 2 compares flexible solutions with the scenario-optimal ones. Flexible solutions are more expensive than deterministic ones, but the cost difference tends to be higher for designs with a lower baseline cost. This indicates that the flexible design approach may require a higher initial investment to enable the WDN to adapt to a range of future scenarios.

6. Conclusions

This work explored the potential of RL in the lifecycle design of WDNs, under deterministic and uncertain conditions. Our experiments, conducted on an adapted NYT benchmark, yielded promising results.

In the deterministic setting, the RL agent had comparable performance to that of the baseline heuristic algorithm and found a cost-effective upgrade strategy for the whole planning horizon despite demand increases and the deterioration of the pipes in the network. Under uncertain conditions, the agent was capable of devising a flexible strategy that could adapt to multiple possible scenarios, an outcome which showcases RL's potential in the lifecycle design of WDNs under deep uncertainty.

While our current results are promising, further validation is needed. Future work will involve applying RL to larger, more realistic networks, incorporating a wider range of future scenarios and exploring how the agent can adapt to emerging information.

Author Contributions: Conceptualization, L.T., C.M., D.S.; methodology, L.T., C.M., D.S.; writing—original draft preparation, L.T.; writing—review and editing, L.T., C.M., D.S.; supervision, C.M., D.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work is a result of the European Research Council (ERC)-funded Water-Futures project (Grant Agreement No. 951424).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Walker, W.E.; Lempert, R.J.; Kwakkel, J.H. Deep Uncertainty. In *Encyclopedia of Operations Research and Management Science*; Springer: Boston, MA, USA, 2013; pp. 395–402.
2. Nagabandi, A.; Clavera, I.; Liu, S.; Fearing, R.S.; Abbeel, P.; Levine, S.; Finn, C. Learning to Adapt in Dynamic, Real-World Environments Through Meta-Reinforcement Learning. *IEEE Trans. Cogn. Dev. Syst.* **2018**, *1*.
3. Tsiami, L.; Makropoulos, C.; Savic, D.A. Reinforcement Learning for Adaptive Water Distribution Network Planning: Exploring its Feasibility and Potential. In Proceedings of the 19th International CCWI Conference, Leicester, UK, 4–7 September 2023.
4. Tsiami, L.; Makropoulos, C.; Savic, D. A review on staged design of water distribution networks. In Proceedings of the 2nd International Joint WDSA/CCWI Conference; Editorial Universitat Politècnica de València, València, Spain, 18–22 July 2022.
5. Schulman, J.; Wolski, F.; Dhariwal, P.; Radford, A.; Klimov, O. Proximal Policy Optimization Algorithms. *arXiv* **2017**, arXiv:1707.06347.
6. Schaake, J.C.; Lai, D. *Linear Programming and Dynamic Programming Application to Water Distribution Network Design*; Report No. 116; Hydrodynamics Laboratory, Department of Civil Engineering, Massachusetts Institute of Technology: Cambridge, MA, USA, 1969.
7. Cunha, M.; Marques, J.; Savić, D. A Flexible Approach for the Reinforcement of Water Networks Using Multi-Criteria Decision Analysis. *Water Resour. Manag.* **2020**, *34*, 4469–4490. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.