Neuro-Symbolic Pump Scheduling for Safe and Cost-efficient Water Distribution Networks

Mahirul Islam[1], Fulong Yao[2], Wanqing Zhao[1], Rafael C. Cardoso[3], and Mengwei Xu[1]

[1] School of Computing, Newcastle University, Newcastle upon Tyne, United Kingdom
{H.M.M.Islam1, Wanqing.Zhao, Mengwei.Xu}@newcastle.ac.uk
[2] School of Computer Science and Informatics, Cardiff University, Cardiff, United Kingdom yaof@cardiff.ac.uk
[3] School of Natural and Computing Sciences, University of Aberdeen, Aberdeen, United Kingdom
rafael.cardoso@abdn.ac.uk

Abstract. The Pump Scheduling Problem is a highly challenging real-world control task in Water Distribution Networks (WDNs) that aims to minimise operational costs while meeting safety requirements (e.g., minimum and maximum allowable tank levels). Latest Deep Reinforcement Learning (DRL) techniques are effective for cost optimisation but can still violate safety constraints at deployment despite explicit safety considerations during training. Furthermore, evolving safety requirements (e.g., due to seasonal considerations) make retraining for minor safety specification changes disproportionately expensive. To address these challenges, we present a neuro-symbolic framework that pairs a pre-trained DRL agent with a symbolic Belief-Desire-Intention (BDI) agent for WDN safety supervision. Our implementation and preliminary empirical results demonstrate improved safety compliance over a DRL-only baseline while maintaining comparable cost performance.

Keywords: Water Distribution Networks · Neuro-Symbolic AI · Deep Reinforcement Learning · BDI Agents.

1 Introduction

Water Distribution Networks (WDNs) are large-scale cyber-physical systems that consist of physical assets (such as pipes, junctions, tanks and pumps) and digital control decisions (e.g., pump schedules) to deliver an essential public service. While their purpose is simple, their operation is non-trivial and energy-intensive, accounting for approximately 2–3% of national electricity consumption in the UK [16], largely because pumping requires significant energy.

As a result, the operation of pumps motivated the *Pump Scheduling Problem* (PSP): an optimisation problem that seeks pump control strategies to minimise operational cost while maintaining hydraulic safety. Recent work has treated

PSP as a sequential decision-making problem and has adopted Deep Reinforcement Learning (DRL) as a practical approach to learn control policies that optimise long-horizon energy cost directly from simulation or operational data, as exemplified in [26].

However, even when safety is explicitly considered during training (e.g., via penalty for water-level violations), DRL policies can still breach safety constraints at deployment. This is because learned policies are inherently stochastic and may still produce occasional safety violations. In addition, safety requirements in WDNs can be time-varying (e.g., across seasons or operator policies), and accommodating small changes in safety specifications can require full retraining or extensive re-tuning, making updates disproportionately costly.

This paper proposes a neuro-symbolic (NeSy) framework that combines data-driven DRL for cost-efficient pump scheduling with a Belief-Desire-Intention (BDI) [5,6] symbolic agent that monitors constraint compliance and intervenes in high-risk situations. We first train a DRL pump scheduler using the latest DRL technique, e.g., Double Duelling Deep Q-Network (D3QN) [24], chosen for its stability in discrete action spaces and its reduced Q-value overestimation. Then, we deploy the successfully trained DRL on a standard WDN benchmark using the EPANET hydraulic simulator[1]. At runtime, the DRL policy proposes pumping control actions, which are checked by a BDI agent implemented in Jason [4] for safety rules and can override unsafe pumping actions before execution in the EPANET simulator. The full framework is implemented[2] to coordinate state exchange and action selection across EPANET, the D3QN controller, and the BDI agent. Preliminary results show a considerable reduction in the frequency of tank-level violations, with only a modest increase in operational cost relative to DRL-only.

The remainder of the paper is organised as follows. Section 2 presents a brief discussion of related work. Section 3 introduces our NeSy framework. Section 4 provides the preliminary empirical results, and Section 5 concludes.

2 Related Work

In this section, we briefly review the related work on the pump scheduling problem in water distribution network optimisation and the existing neuro-symbolic (NeSy) approaches, especially those using the BDI paradigm.

The Pump Scheduling Problem (PSP) has been well-studied, with early formulations dating back at least to 1989 [21]. A wide range of techniques is developed to minimise cost while meeting hydraulic constraints, including linear programming [20,13], non-linear programming [22,25], and metaheuristics [26,7,12].

More recently, PSP has been treated as a sequential decision-making problem [10], which has motivated the use of DRL to learn control policies directly from simulated or real operational data [19]. For example, the key work in [9] reports DRL-based pump scheduling that achieves higher cost efficiency than

[1] `https://www.epa.gov/water-research/epanet/` (Accessed: 27/04/2026)
[2] `https://github.com/Mahir-Islam/hydro_mas` (Accessed: 27/04/2026)

conventional optimisation techniques in [18,2]. That said, cost-effective DRL-based controllers do not guarantee constraint satisfaction at deployment, where occasional safety violations can still occur.

Meanwhile, the Neuro-Symbolic (NeSy) approach integrates neural models with symbolic reasoning techniques to combine data-driven learning with explicit rule-based deliberation [15]. NeSy architectures can be represented using dual-process theory [14,3]: the neural component (System 1) typically performs rapid inference under routine conditions, while the symbolic component (System 2) enforces constraints when higher levels of scrutiny are required. There have been NeSy architectures developed for efficient and safety-critical domains by combining a sub-symbolic component with a symbolic partner. For example, one of the closest works to ours is [1], which uses a symbolic BDI agent to supplement machine learning models in their decision-making to ensure driving safety in simulated autonomous driving. However, while they employ an orchestrator to mediate decisions between the two systems, we instead centralise control in the BDI agent for improved safety assurances.

3 DRL–BDI Framework

Our proposed NeSy approach balances two competing objectives: maintaining safe tank levels and minimising the energy cost of pump operation. Our framework couples three components: (i) a DRL controller that proposes cost-efficient pump actions, (ii) a symbolic BDI supervisor that enforces hydraulic safety, and (iii) a hydraulic simulation environment that executes actions and returns the evolving WDN state. Figure 1 overviews the resulting DRL–BDI control loop.

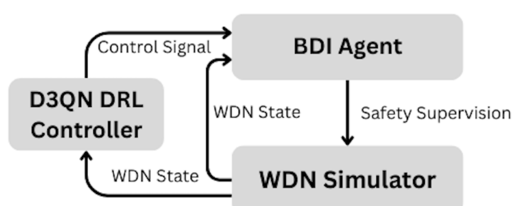

Fig. 1. DRL–BDI framework applied to WDN.

3.1 Water Distribution Network Environment

We apply our framework to *D-Town* [17], a widely used benchmark WDN that approximates the water supply for a small city, as depicted in Fig. 2. D-Town comprises 400 junctions, 11 pumps, and 7 tanks arranged into 5 demand zones. Each demand zone contains a pump group (a cluster of pumps operated as a single unit), a primary storage tank, and a subset of junction demands. Hydraulic dynamics are simulated using EPANET, which advances the network state at an hourly timestep. At each timestep, the simulator (i) receives pump actions, (ii) performs the hydraulic computation for the next hour, and (iii) returns the updated WDN state, including tank levels and energy usage.

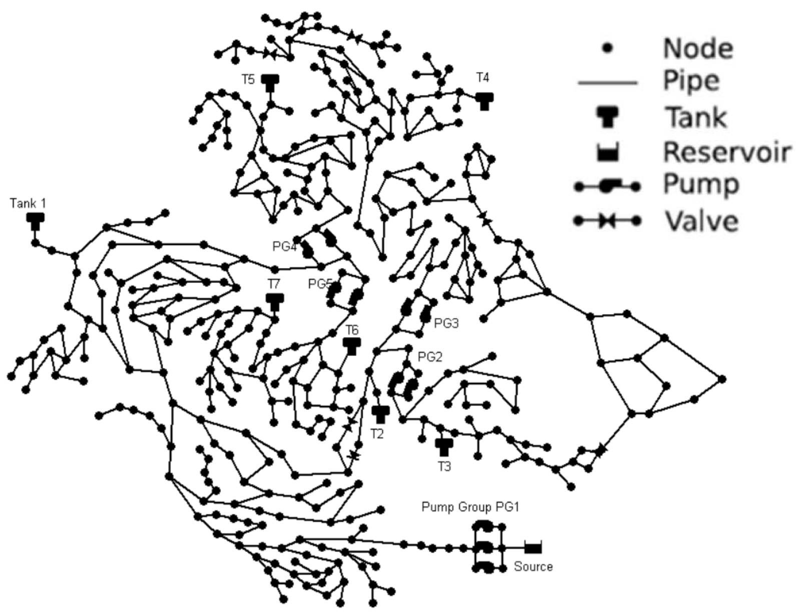

Fig. 2. Topological graph of the D-Town benchmark WDN.

3.2 DRL Agent

The DRL agent uses a Double Duelling Deep Q-Network (D3QN) [24] that proposes pump-group actions to minimise energy cost while keeping tank levels within safe operating ranges. It employs a holistic data-driven control scheme for sequential decision-making problems. For improved learning efficiency and accuracy in reward estimation, D3QN consists of a Q-Network and a Target Network, both of which contain a state-value function $V(s)$ and an advantage function $A(s,a)$. Here, the target network is a copy of the Q network that is used to calculate future rewards, and it is only updated every few timestep. The Q-value determining the expected reward of a control action is thereby calculated as: $Q(s,a) = V(s) + A(s,a) - mean(A(s,a))$; full details can be found in [23].

At each timestep, the agent observes a state vector s summarising recent demand/price history and current tank levels from the EPANET simulator, then returns an action a that optimises the control of each pump group. We now briefly describe how we train a DRL agent using D3QN for the D-Town scenario.

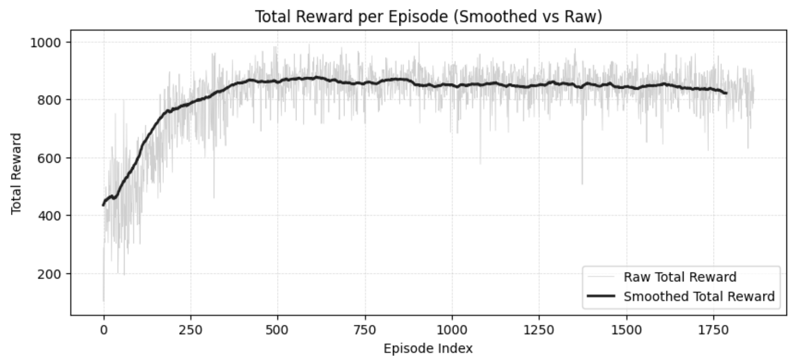

Fig. 3. Training reward over episodes for D3QN controller. Convergence occurs around episode 600–700.

There are five pump groups in D-Town, each with two operational settings: on (1) or off (0). An action is therefore a 5-bit activation vector. Following common practice for discrete Q-learning, we encode actions as integers in $[0, 31]$ and decode them to binary vectors, e.g., $31_{10} = 11111_2 \mapsto [1, 1, 1, 1, 1]$. All pumps within a group are operated together rather than independently. Meanwhile, the state comprises demand, price, and tank levels. We use a context window of 6 timesteps (6 hours)[3] to expose the short-term temporal structure to the Q-network. All tanks (including those indirectly controlled by a pump) take into consideration the violation of constraints to provide a complete hydraulic context. Consequently, the training objective is a weighted reward combining cost and safety where $w_1, w_2 \in [0, 1]$ and $w_1 + w_2 = 1$:

$$r_k = w_1 \cdot \text{cost\_value}(k) + w_2 \cdot \text{safety\_value}(k) \tag{1}$$

Safety is assessed by normalised tank levels $[0, 1]$: the *safe bounds* are $[0.2, 0.8]$, while the simulator itself enforces *hard bounds* $[0.1, 0.9]$ to avoid infeasible hydraulics. Deviations from safe bounds incur penalties via a scaled trapezoidal membership function with parameters $\{0, 0.2, 0.8, 1\}$ (with minimum value -1). Cost is computed as total pump energy usage multiplied by the electricity price, and normalised using an inverse-logit transformation to stabilise learning.

We train the D3QN controller on historical demand and price profiles until the return stabilises and the policy consistently achieves low operating cost with few safety violations. Figure 3 confirms the reward convergence during training.

3.3 BDI Agent

The symbolic agent is instantiated as a BDI agent implemented in *Jason*, acting as a symbolic safety supervisor over the DRL controller. BDI agents achieve

[3] Optimal value found after hyperparameter tuning during training.

this by double-checking the control signals from the DRL agent before passing to the EPANET simulator. In our setting, beliefs are grounded in the current WDN state returned by EPANET, including: tank levels, demand levels, and pump-group states. The agent's top priority is to maintain all tanks within safe bounds. In our framework, the BDI agent operationalises this by encoding domain safety knowledge as a rule base structured into a small hierarchy of increasingly proactive decisions. It first applies a *corrective* action that reacts to explicit constraint violations and issues immediate remedial actions. If no violation is present, it applies *anticipatory* reasoning [11] that examines near-term trends and intervenes when projected trajectories indicate an impending breach of the anticipatory safety range. If neither of the two above is applicable, the agent will reason over the zone demand to improve distribution efficiency; for example, the agent may decide to pump more when the demand is high. Finally, a *fallback* action accepts and executes the DRL-proposed action when no symbolic rule applies, preserving the learned controller's cost efficiency whenever possible.

3.4 DRL–BDI Implementation

We have implemented the overall framework, and the source code is publicly available online[4]. In detail, the Python-based DRL controller, the Jason BDI agent, and the EPANET simulator are integrated in a closed loop. At each timestep, the simulator provides the current WDN state to the DRL agent. The DRL controller proposes an action, which will also be passed to BDI agents as a proposed action, by evaluating its Q-network on the corresponding state vector. Meanwhile, the EPANET simulator also passes the network state to BDI agents as the knowledge. The BDI agent then deliberates and either (i) overrides the proposal to enforce safety/anticipation rules or (ii) accepts it via the fallback mechanism. The selected action is finally applied to the EPANET simulator to advance the network to the next timestep. In our current implementation, the DRL–BDI bridge uses a lightweight file-based inter-process communication mechanism to exchange state and actions across the Python and Java runtimes.

4 Evaluation

The neural network has a batch size of 64, discount factor of 0.99, learning rate of 0.001, and an initial exploration rate of 0.9, decaying by 0.5% every timestep, until it reaches a floor of 0.1.

We evaluate our NeSy framework by comparing it against a D3QN-only baseline (henceforth referred to as DRL). Specifically, we assess (i) the rate of constraint violations (how often tank levels leave the permitted range), (ii) the severity of safety deviations (how far and for how long tank levels drift outside the safe bounds), and (iii) operational cost. To evaluate, we have run the full 168-timestep (constituting 168 hours) D-Town simulation for each system. We then collect time-series data for cost, constraint violations, and violation severity.

[4] `https://github.com/Mahir-Islam/hydro_mas` (Accessed: 27/04/2026)

Table 1. Tank Level Constraint Compliance over 168 hours (Tank 2 excluded)

Metric	Tank 1	Tank 3	Tank 4	Tank 5	Tank 6	Tank 7	Overall
DRL (% Compliance)	84	64	61	80	98	69	76
Hydro-MAS (% Compliance)	97	94	62	92	83	81	85

Table 1 shows that the NeSy controller improves overall constraint compliance from 76% to 85%, representing a 9 percentage-point increase over DRL-only control. These improvements are consistent across most tanks, except tanks 4 and 6. Tank 6 exhibits reduced performance under our approach because it is out of our direct control. The network pump ruleset only considers Tank 7 levels (Tanks 6 and 7 are affected by the same pump group). Both controllers show lower compliance for Tank 4, since rapid level fluctuations with lag time makes control difficult. Tank 2 is not included, since it operates independently of the action space.

Table 2. Wilcoxon Signed-Rank Test Results for Safety Metrics ($n = 168 \times 6 = 1008$).

Metric	Mean (DRL)	Mean (NeSy)	Wilcoxon p-value
Deviation	0.1911	0.1682	0.00027
Violation Severity	0.0814	0.0553	5.0e-08
Violation Occurrence	0.2402	0.1520	3.2e-07

Table 2 shows the severity of safety deviations using three derived metrics computed over all tanks (excluding Tank 2), yielding $n = 168 \times 6 = 1008$ observations per controller. These three metrics can be determined from the tank levels: the *deviation* is defined as the deviation from the middle tank level (0.5); the *violation severity* is the positive excess beyond the safe interval [0.2, 0.8] with negative values set to zero; finally, *violation occurrence* is a binary indicator equal to 1 if the violation severity is greater than 0, otherwise 0. For each metric, a Wilcoxon signed-rank test [8] was conducted to compare paired samples without assuming normality or homogeneity of variances. It is a statistical test used to determine whether two samples differ significantly. This Wilcoxon signed-rank test shows that the safety measures of the NeSy approach differ significantly from those in the DRL-only setting.

Finally, we examine the operational cost of these two approaches. Operational cost is measured using the normalised cost value over 168 hours. Fig. 4 shows the comparison of the cost. Overall, there is a modest and statistically significant increase in the cost. The DRL-only controller achieves an average normalised cost of 0.168, whereas our NeSy approach achieves the average normalised cost of 0.218.

In summary, our NeSy DRL–BDI framework improves WDN safety metrics (reduced deviation, reduced violation prevalence, and reduced violation severity) while incurring a modest increase in operational cost relative to DRL-only.

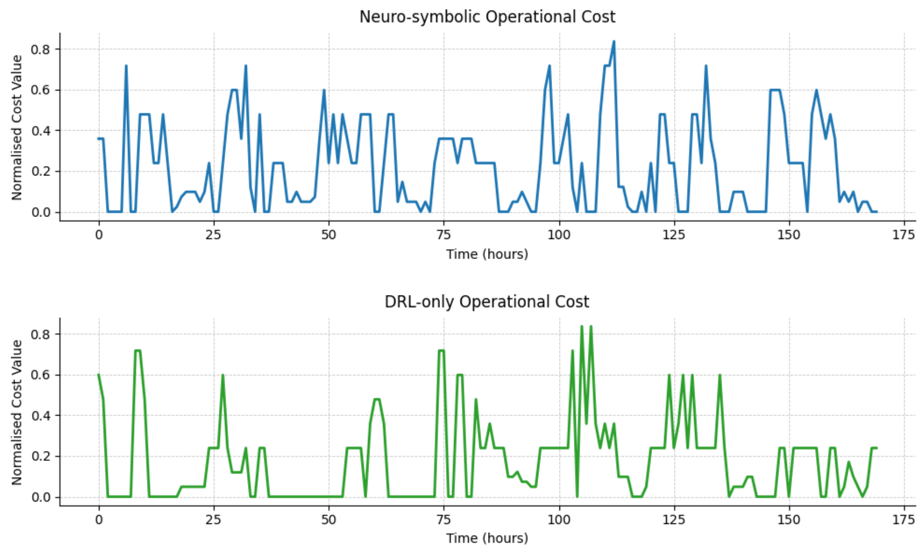

Fig. 4. Cost of operation over time.

5 Conclusion and Future Work

This paper presented a neuro-symbolic DRL–BDI framework for pump scheduling in WDNs that addresses a practical tension: lowering operational energy costs while maintaining safe tank levels at runtime. The key idea is to retain the strong cost performance of a successfully trained D3QN controller while placing it under symbolic supervision from a BDI agent that can veto or amend unsafe control proposals. On the standard D-Town benchmark, our preliminary results show that this hybrid supervision improves safety compared to a DRL-only baseline, reducing the frequency and severity of safe-bound deviations while keeping operational cost broadly comparable.

Several directions follow naturally from these findings. First, the evaluation here is intentionally preliminary and should be strengthened with longer horizons and broader operating conditions (e.g., diverse demand/price profiles and stress scenarios). Second, we will extend the supervisor beyond tank-level bounds to additional operational objectives such as pump switching frequency (wear), pressure-related constraints, and water quality considerations. We also plan to improve the implementation with a more efficient messaging interface to scale to longer or higher-frequency simulations. Finally, a major avenue for future work is to improve the symbolic supervisor itself. In particular, we will refine and systematically tune its rule base and intervention strategy to further reduce violation rates while preserving the cost-efficiency of the underlying DRL policy.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Al Shukairi, H., Cardoso, R.C.: ML-MAS: A hybrid AI framework for self-driving vehicles. In: Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems. p. 1191–1199. AAMAS '23, International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC (2023)
2. Baltar, A.M., Fontane, D.G.: Use of multiobjective particle swarm optimization in water resources management. Journal of water resources planning and management **134**(3), 257–265 (2008)
3. Booch, G., Fabiano, F., Horesh, L., Kate, K., Lenchner, J., Linck, N., Loreggia, A., Murgesan, K., Mattei, N., Rossi, F., Srivastava, B.: Thinking fast and slow in ai. Proceedings of the AAAI Conference on Artificial Intelligence **35**(17), 15042–15046 (May 2021). `https://doi.org/10.1609/aaai.v35i17.17765`
4. Bordini, R.H., Wooldridge, M., Hübner, J.F.: Programming Multi-Agent Systems in AgentSpeak using Jason. John Wiley & Sons, Chichester, UK (2007)
5. Bratman, M.: Intention, plans, and practical reason (1987)
6. Cardoso, R.C., Ferrando, A.: A review of agent-based programming for multi-agent systems. Computers **10**(2), 16 (2021). `https://doi.org/10.3390/computers10020016`
7. da Conceição Cunha, M., Sousa, J.: Water distribution network design optimization: Simulated annealing approach. Journal of Water Resources Planning and Management **125**(4), 215–221 (1999). `https://doi.org/10.1061/(ASCE)0733-9496(1999)125:4(215)`
8. Conover, W.J.: Practical nonparametric statistics (3rd ed.). John Wiley & Sons, Inc. (1999)
9. Hajgató, G., Paál, G., Gyires-Tóth, B.: Deep reinforcement learning for real-time optimization of pumps in water distribution systems. Journal of Water Resources Planning and Management **146**(11), 04020079 (2020). `https://doi.org/10.1061/(ASCE)WR.1943-5452.0001287`
10. Hu, S., Gao, J., Zhong, D., Wu, R., Liu, L.: Real-time scheduling of pumps in water distribution systems based on exploration-enhanced deep reinforcement learning. Systems **11**(2) (2023). `https://doi.org/10.3390/systems11020056`
11. Hübner, J.F., Burattini, S., Ricci, A., Mayer, S.: Reflexive anticipatory reasoning by BDI agents. Auton. Agents Multi Agent Syst. **39**(1), 7 (2025). `https://doi.org/10.1007/S10458-025-09687-8`, `https://doi.org/10.1007/s10458-025-09687-8`
12. Jandaghi, H., Haghgoo, A., Tavakoli Moghadam, M.: Optimization of water distribution networks using simulated annealing (SA) with variable neighbourhood search (VNS): a case study in the city of malard. Water Practice and Technology **19**(12), 4958–4986 (12 2024). `https://doi.org/10.2166/wpt.2024.295`
13. Jowitt, P.W., Germanopoulos, G.: Optimal pump scheduling in water-supply networks. Journal of Water Resources Planning and Management **118**(4), 406–422 (1992). `https://doi.org/10.1061/(ASCE)0733-9496(1992)118:4(406)`
14. Kahneman, D.: Thinking, fast and slow. Farrar, Straus and Giroux, New York (2011)
15. Kautz, H.: The third ai summer: Aaai robert s. engelmore memorial lecture. AI Magazine **43**(1), 105–125 (Mar 2022). `https://doi.org/10.1002/aaai.12036`
16. Majid, A., Cardenes, I., Zorn, C., Russell, T., Colquhoun, K., Bañares-Alcantara, R., Hall, J.W.: An analysis of electricity consumption patterns in the water and wastewater sectors in south east england, uk. Water **12**(1) (2020). `https://doi.org/10.3390/w12010225`

17. Marchi, A., Salomons, E., Ostfeld, A., Kapelan, Z., Simpson, A.R., Zecchin, A.C., Maier, H.R., Wu, Z.Y., Elsayed, S.M., Song, Y., Walski, T., Stokes, C., Wu, W., Dandy, G.C., Alvisi, S., Creaco, E., Franchini, M., Saldarriaga, J., Páez, D., Hernández, D., Bohórquez, J., Bent, R., Coffrin, C., Judi, D., McPherson, T., van Hentenryck, P., Matos, J.P., Monteiro, A.J., Matias, N., Yoo, D.G., Lee, H.M., Kim, J.H., Iglesias-Rey, P.L., Martínez-Solano, F.J., Mora-Meliá, D., Ribelles-Aguilar, J.V., Guidolin, M., Fu, G., Reed, P., Wang, Q., Liu, H., McClymont, K., Johns, M., Keedwell, E., Kandiah, V., Jasper, M.N., Drake, K., Shafiee, E., Barandouzi, M.A., Berglund, A.D., Brill, D., Mahinthakumar, G., Ranjithan, R., Zechman, E.M., Morley, M.S., Tricarico, C., de Marinis, G., Tolson, B.A., Khedr, A., Asadzadeh, M.: Battle of the water networks ii. Journal of Water Resources Planning and Management **140**(7), 04014009 (2014). https://doi.org/10.1061/(ASCE)WR.1943-5452.0000378
18. Odan, F.K., Ribeiro Reis, L.F., Kapelan, Z.: Real-time multiobjective optimization of operation of water supply systems. Journal of Water Resources Planning and Management **141**(9), 04015011 (2015). https://doi.org/10.1061/(ASCE)WR.1943-5452.0000515
19. Pei, S., Hoang, L., Fu, G., Butler, D.: Real-time pump scheduling in water distribution networks using deep reinforcement learning. Journal of Water Resources Planning and Management **151**(6), 04025012 (2025). https://doi.org/10.1061/JWRMD5.WRENG-6476
20. Shao, Y., Zhou, X., Yu, T., Zhang, T., Chu, S.: Pump scheduling optimization in water distribution system based on mixed integer linear programming. European Journal of Operational Research **313**(3), 1140–1151 (2024). https://doi.org/https://doi.org/10.1016/j.ejor.2023.08.055
21. Tarquin, A.J., Dowdy, J.: Optimal pump operation in water distribution. Journal of Hydraulic Engineering **115**(2), 158–168 (1989). https://doi.org/10.1061/(ASCE)0733-9429(1989)115:2(158)
22. Vieira, B.S., Mayerle, S.F., Campos, L.M., Coelho, L.C.: Optimizing drinking water distribution system operations. European Journal of Operational Research **280**(3), 1035–1050 (2020). https://doi.org/https://doi.org/10.1016/j.ejor.2019.07.060
23. Wang, Z., de Freitas, N., Lanctot, M.: Dueling network architectures for deep reinforcement learning. CoRR **abs/1511.06581** (2015), http://arxiv.org/abs/1511.06581
24. Yao, F., Zhao, W., Forshaw, M., Song, Y.: A holistic power optimization approach for microgrid control based on deep reinforcement learning. Neurocomputing **654**, 131375 (2025). https://doi.org/https://doi.org/10.1016/j.neucom.2025.131375
25. Yu, G., Powell, R.S., Sterling, M.J.H.: Optimized pump scheduling in water distribution systems. J. Optim. Theory Appl. **83**(3), 463–488 (Dec 1994). https://doi.org/10.1007/BF02207638
26. Zhao, W., Beach, T.H., Rezgui, Y.: Optimization of potable water distribution and wastewater collection networks: A systematic review and future research directions. IEEE Transactions on Systems, Man, and Cybernetics: Systems **46**(5), 659–681 (2016). https://doi.org/10.1109/TSMC.2015.2461188