ICA2025OSLO
14th IWA Specialised Conference on
Instrumentation, Control and Automation

# Model Predictive Reinforcement Learning Control based on Liquid Neural Networks for Global Plant-wide Control of a WRRF

C. C. Gómez Cortés*,**, K. Solon*,**, P. Ramin***, X. Flores-Alsina***, J. K. Huusom***, M. Külahci****

E. Torfs*,**,*****

* BIOMATH, Department of Data Analysis and Mathematical Modelling, Faculty of Bioscience Engineering, Ghent University, Coupure links 653, 9000 Ghent, Belgium

** Centre for Advanced Process Technology for Urban REsource recovery (CAPTURE), Frieda Saeysstraat 1, 9052 Ghent, Belgium

*** Process and Systems Engineering Centre (PROSYS), Department of Chemical and Biochemical Engineering, Technical University of Denmark, Søltofts Plads, Building 228 A, 2800 Kgs. Lyngby, Denmark

**** Statistics and Data Analysis, Department of Applied Mathematics and Computer Science, Technical University of Denmark, Richard Petersens Plads, Building 324, 2800 Kgs. Lyngby, Denmark

***** modelEAU, Department of Civil Engineering and Water Engineering, Laval University, 1065 avenue de la Médecine, Québec G1V 0A6, QC, Canada

cristiancamilo.gomezcortes@ugent.be, kimberly.solon@ugent.be, pear@kt.dtu.dk, xfa@kt.dtu.dk, jkh@kt.dtu.dk, muku@dtu.dk, elena.torfs.1@ulaval.ca

**Abstract:** This study implemented Liquid Neural Networks (LNN) within a Reinforcement Learning (RL) control framework to optimize Water Resource Recovery Facilities (WRRFs). Compared to PID controllers, RL control offered better response to nonlinear disturbances and dynamic scenarios, such as rain events and fluctuating energy prices. The tested scenario involving increased peak influent flowrates with energy price variation demonstrated RL's ability to maintain improved effluent quality, by mitigating effluent nitrogen peaks, while achieving comparable energy consumption to PID control. The findings highlight the potential of RL-based control for global plant-wide control of WRRF operations.

**Keywords**: Global control; Reinforcement Learning; Liquid Neural Networks; WRRF

## INTRODUCTION

Water resource recovery facilities (WRRFs) are vital for pollution reduction but face multiple operational challenges such as high energy and resource use, greenhouse gas emissions, and increasingly stringent effluent quality limits, to name a few. Control and optimization within WRRFs are particularly complex due to the high nonlinearity inherent in the biological and dynamic processes (Olsson, G., 2012).

Conventional Proportional-Integral-Derivative (PID) controllers are effective under normal operational conditions but struggle with nonlinear disturbances such as extreme weather. These controllers typically maintain constant setpoints, which are not always optimal for changing conditions or achieving specific operational goals. Moreover, they are mainly implemented locally making it challenging to simultaneously address multiple, often conflicting, operational goals.

Addressing these limitations necessitates a global control methodology that can adapt under normal and extreme scenarios. Unlike with PID control, Reinforcement Learning (RL) control dynamically adjusts setpoints in response to changing conditions, enabling greater optimization while addressing multiple WRRF challenges.

Although machine learning methodologies such as Deep and Multi-agent RL have been applied in this context, the large number of parameters required limit their interpretability and transferability (Hernández-del-Olmo et al., 2023). To address these limitations, this study implements Liquid Neural Networks (LNNs) (Hasani et al., 2021) within the RL control framework. LNNs are particularly suited for temporal and causal reasoning due to their unique architecture based on continuous differential equations. Unlike fully connected neural networks, LNNs feature a sparse architecture, reducing computational complexity and improving interpretability. These characteristics enable LNNs to incorporate causality and time-dependent

dynamics into the control strategy, making them well-suited for optimizing WRRF operations under both normal and extreme scenarios.

## METHODOLOGY

**System under study.** The WRRF is based on the Benchmark Simulation Model No.2 plant (Jeppsson et al., 2007) with minor modifications (see **Figure 1**). The water line consists of a primary clarifier, an activated sludge unit with a Modified Ludzak-Ettinger (MLE) configuration, and a secondary clarifier. The sludge line includes an anaerobic digester (AD) preceded by thermal hydrolysis pre-treatment (THP) to enhance sludge degradation and biogas production. Variable storage tanks are included to manage external sludge and return reject water. A thickener and a dewatering unit are also included for handling sludge before and after AD, respectively. The baseline plant utilizes PID controllers tuned with the Ziegler-Nichols method to establish a benchmark for evaluating the RL control performance. The controllers are detailed in **Table 1**.

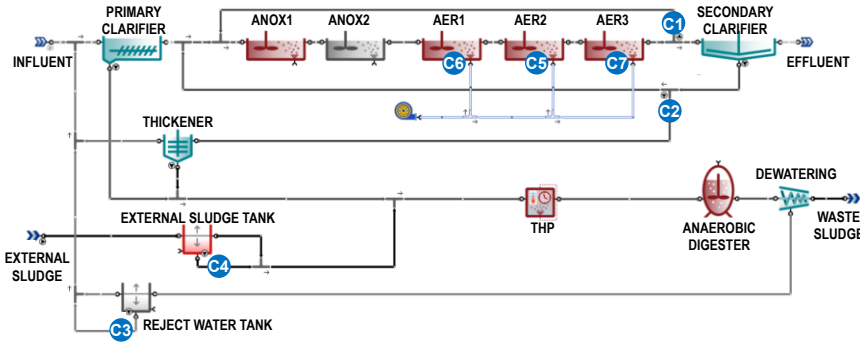



**Figure 1** Plant-wide WRRF under study

## Reinforcement Learning (RL) control

***RL Training.*** The training is done offline, where the RL agent learns decision-making by interacting with the environment (see **Figure 2**) to maximize a reward function. The agent employs a LNN, a variant of recurrent neural network characterized by continuous differential equations for spatiotemporal decision-making (Hasani, et al., 2020; Croll, et al., 2023). A specialized set of neuron compartments was engineered encompassing sensory processing, control, and motor output to optimize the network's design for sparse efficiency. Once trained, the RL agent dynamically adjusts the setpoints of the PID controllers.

**Table 1.** Implemented controllers in the WRRF baseline model under study

| | Default Control Settings | | | RL Adjusted Parameter |
|---|---|---|---|---|
| | Type | Control Variable | Manipulated Variable | |
| 1 | PID | Nitrate ($NO_x$) in AER3 | Internal recirculation pump flow | $NO_x$ setpoint in AER3 |
| 2 | PID | TSS in AER3 | Sludge recirculation pump flow | TSS setpoint in AER3 |
| 3 | PID | Volume of reject water | Pumped flow out of reject water tank | $Volume_{reject\ water\ tank}$ |
| 4 | PID | Volume of external sludge | Pumped flow out of external sludge tank | $Volume_{external\ sludge\ tank}$ |
| 5 | PID (Cascade) | $NH_x$ in AER2 | DO setpoint in AER2 | $NH_x$ setpoint in AER2 |
| | | DO setpoint in AER2 | Airflow in AER2 | |
| 6 | Ratio | Airflow in AER2 | Airflow ratio in AER1 (i.e., 150% of AER2) | Airflow ratio in AER1 |
| 7 | Ratio | Airflow in AER2 | Airflow ratio in AER3 (i.e., 50% of AER2) | Airflow ratio in AER3 |

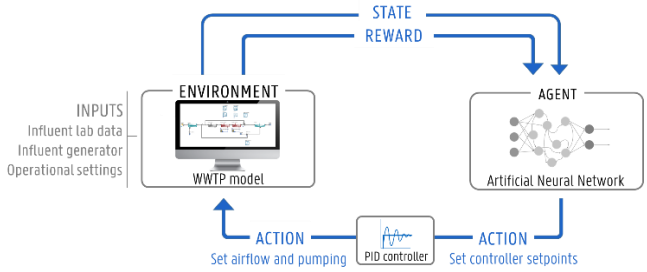



**Figure 2** Schematic of RL training

***Reward function***. To simplify the demonstration of the LNN-based RL control, the comparison is focused on two key performance criteria: effluent quality, specifically total nitrogen (TN) and the total energy consumption (due to aeration, pumping, etc.). Thus, the reward function in **Equation 1** incorporates these two metrics.

The reward function follows a hierarchical structure, prioritizing the control of excess effluent TN, with managing energy consumption as a secondary focus. It is designed to address priorities sequentially. When effluent TN levels exceed the reference, the priority shifts entirely to minimizing deviations from the reference, excluding energy considerations to prioritize reducing effluent TN. If effluent TN is under control, the priority moves to reducing energy consumption when it increases above the reference value. Under stable conditions, the function balances both objectives, simultaneously optimizing TN and energy reduction.

$$r_t = \begin{cases} \text{if } C_{\text{eff TN}_{t\to 24h}} \geq C_{\text{eff}_{\text{TN}_{\text{Limit}}}},\ 10^2 \cdot \left( \sum_{i=t}^{t+24h} \left( \frac{C_{\text{eff TN}_{\text{Target}}} - C_{\text{eff TN}_t}}{C_{\text{eff TN}_{\text{Target}}}} \right)^2 \right) \\ \text{if } E_{\text{Total}_{t\to 24h}} \geq E_{\text{Total}_{\text{Year_max}}},\ 10^1 \cdot \left( \sum_{i=t}^{t+24h} \left( \frac{E_{\text{Total}_t}}{E_{\text{Total}_{\text{Year_average}}}} \right)^2 \right) \\ \text{else, } 10^0 \cdot \left( \sum_{i=t}^{t+24h} \left( \frac{C_{\text{eff TN}_{\text{Target}}} - C_{\text{eff TN}_t}}{C_{\text{eff TN}_{\text{Target}}}} \right)^2 + \sum_{i=t}^{t+24h} \left( \frac{E_{\text{Total}_t}}{E_{\text{Total}_{\text{Year_average}}}} \right)^2 \right) \end{cases} \qquad \textbf{Equation 1}$$

Where: $C_{\text{eff}_{\text{TN}_{\text{Limit}}}} = 12.5\ \text{gN.m}^{-3}$, $C_{\text{eff TN}_{\text{Target}}} = 10.0\ \text{gN.m}^{-3}$, $E_{\text{Total}_{\text{Year_average}}} = 420\text{kWh}$, $E_{\text{Total}_{\text{Year_max}}} = 520\ \text{kWh}$

***Scenarios for training and testing the RL-based controller.*** To train the RL controller, scenarios of increased peak influent flowrate and varying energy price were simulated. During rain events, peak influent flowrate was stochastically generated to increase 1.5-2 times the baseline levels, with 2-3 random events occurring over a 30-day period. Influent concentrations remain constant. Price variation is introduced through changes in energy costs based on energy market dynamics (**Figure 3b**). The tested scenario reflects the influent flowrate profile shown in **Figure 3a**.

***Implementation details.*** SUMO (©Dynamita), a mechanistic modelling platform, simulated the WRRF processes, while Python handled RL training with the LNN model interfaced to SUMO. Finally, the trained agent was deployed in SUMO to simulate the test scenario.

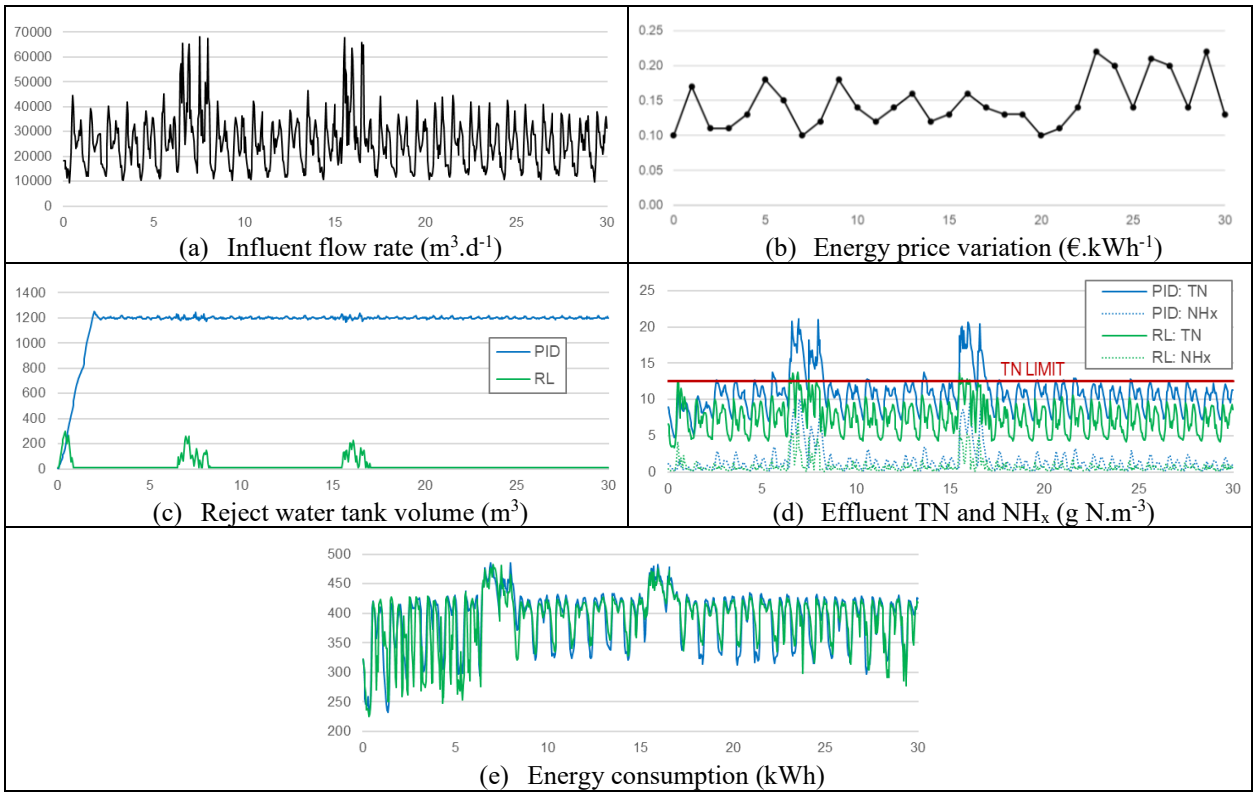
## RESULTS AND DISCUSSIONS

The results for the scenario of increased peak influent flowrate and varying energy price are presented in **Figure 3**. Under PID control, the liquid volume in the reject water tank was maintained at its setpoint of 1200 m³, reflecting the system's compliance to fixed operational settings and therefore limits the system's ability to adapt dynamically to changing process conditions. In contrast, RL demonstrated significant improvements in tank management by maintaining minimal volumes, below 300 $m^3$, during normal operations and strategically filling the tank (**Figure 3c**) coinciding with the major flow peaks seen around days 7-8 and 16-17 (**Figure 3a**).

For effluent quality, PID control showed TN concentration spikes substantially exceeding the limit of 12.5 g $N.m^{-3}$ during high flow events. On the other hand, RL mitigated peaks in effluent TN and $NH_x$ concentrations while consistently keeping TN levels below the set limit (**Figure 3d**).

The energy price variations (**Figure 3b**) show daily fluctuations between 0.10-0.22 €.$kWh^{-1}$. Despite these variations, energy consumption remains relatively comparable between PID and RL-based controls (**Figure 3e**). However, RL demonstrated a clear advantage by maintaining improved effluent TN quality and more effective reject water tank storage management, showcasing its ability to comply with effluent TN limit, while maintaining energy consumption comparable to PID, even under dynamic high-flow conditions.

An important aspect of the implemented control strategy is the integration of the mechanistic model's 24-hour ahead predictions into the RL reward function (**Equation 1**) enabling Model Predictive Control (MPC) capabilities. This approach allows the controller to anticipate for upcoming disturbances rather than merely reacting to current conditions. This is evident in the results where the RL controller pre-emptively adjusts tank levels before actual flow peaks arrive and maintains stable effluent TN quality despite significant influent flow variations.




(a) Influent flow rate ($m^3.d^{-1}$)
(b) Energy price variation ($€.kWh^{-1}$)
(c) Reject water tank volume ($m^3$)
(d) Effluent TN and $NH_x$ ($g\ N.m^{-3}$)
(e) Energy consumption (kWh)

**Figure 3** Results of the WRRF performance. All horizontal axes represent days.

## CONCLUSIONS

- RL control improves operational adaptability by mitigating effluent TN and $NH_x$ peaks. This is done dynamically by adjusting control parameters, overcoming the limitations of conventional PID control in handling nonlinear disturbance, i.e. maintaining fixed set-points.
- The findings suggest the potential of integrating RL-based control within digital twins, with opportunities to expand reward functions to address multiple objectives.

The full manuscript will provide additional performance metrics, along with results of RL control considering other pollutant concentrations, greenhouse gas emissions, etc. In addition, comparison of LNN with other deep neural network-based control (i.e., RNN) will be included, as well as analyses of additional operational scenarios.

## REFERENCES

Chen, K., Wang, H., Valverde-Pérez, B., Zhai, S., Vezzaro, L., & Wang, A. (2021). Optimal control towards sustainable wastewater treatment plants based on multi-agent reinforcement learning. Chemosphere, 279, 130498.

Croll, H. C., Ikuma, K., Ong, S. K., & Sarkar, S. (2023). Reinforcement learning applied to wastewater treatment process control optimization: Approaches, challenges, and path forward. Crit. Rev. Environ. Sci. Technol., 53(20), 1775-1794.

Hasani, R., Lechner, M., Amini, A., Rus, D., & Grosu, R. (2020). Liquid Time-constant Networks. arXiv:2006.04439 [cs.LG].

Hernández-del-Olmo, F., Gaudioso, E., Duro, N., Dormido, R., & Gorrotxategi, M. (2023). Advanced Control by Reinforcement Learning for Wastewater Treatment Plants: A Comparison with Traditional Approaches. Appl. Sci., 13(8), 4752.

Jeppsson, U., Pons, M.N., Nopens, I., Alex, J., Copp, J. B., Gernaey, K.V., Rosen, C., Steyer, J.P., & Vanrolleghem, P.A. (2007). Benchmark Simulation Model No.2: General protocol and exploratory case studies. Water Sci. Technol., 56(8), 67-78.

Olsson, G. (2012). ICA and me – A subjective review. Water Res., 46(6), 1585-1624.