DARROW

Technical Results

Driving the **Future** of **Water Resource Recovery Facilities** through **Data Intelligence**

TABLE OF CONTENTS

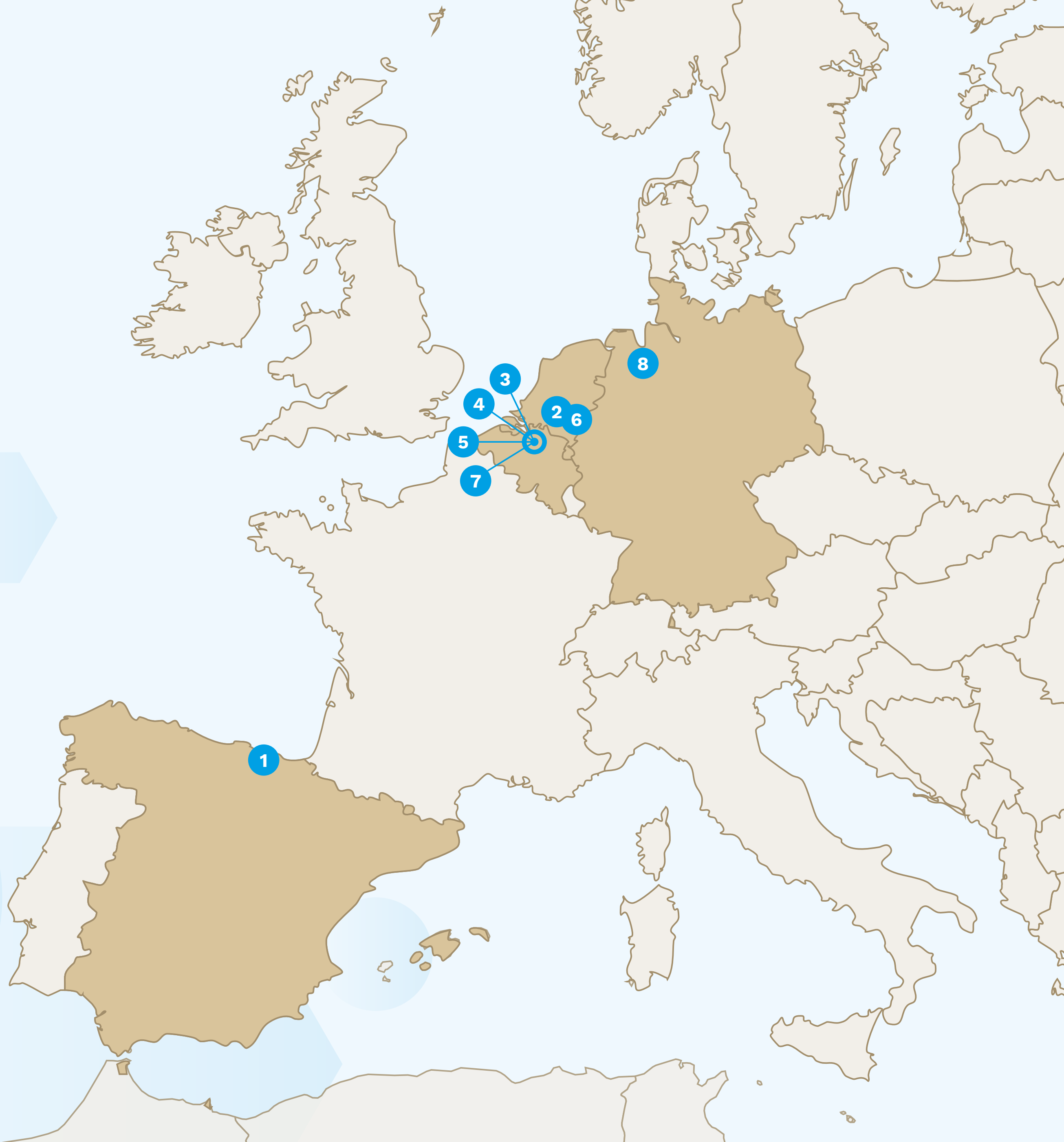
Picture credits

Cover and pages 9, 19, 24, 27, 29, 31 - ESCI; Page 15, 22, 23 - Ceit; Page 18 - IMEC;

About the project

Driving the future of water resource recovery facilities through data intelligence.

DARROW aims to build a data-driven AI solution to optimise resource recovery from wastewater by making wastewater treatment plants more autonomous and energy efficient.



ES / RTO

Coordinator. Lead on ROM models, digital twins, software sensors, project management.

1

NL / IND

End-user and operator of the Tilburg demonstration site. KPI definition and pilot assessment.

2

BE / UNI

Lead on mechanistic modelling, RL training environment, dynamic simulators, operator training (MOOC).

3

NL / SME

Data augmentation services, N_2O risk modelling, AI/ML methodology, integration support.

4

BE / RTO

Lead on RL model development, Explainable AI methods, human-technology interaction and social impact.

5

NL / SME

Lead on software integration, deployment (Aquasuite platform), and commercial exploitation.

6

BE / RTO

Lead on data augmentation and data enrichment AI/ML models.

7

DE / NPO

Lead on communication and dissemination.

8

01

The challenge: From wastewater treatment to resource recovery

Wastewater treatment plants are undergoing a profound transformation. Once designed primarily to protect public health and the environment by treating wastewater before discharge, they are increasingly expected to recover valuable resources, reduce greenhouse gas emissions, improve energy efficiency and contribute to a circular economy.

Meeting these expectations requires more than incremental improvements. It calls for smarter, data-driven operation that helps utilities optimise increasingly complex processes while maintaining reliable and compliant treatment. This is the challenge DARROW set out to address.

1.1 Unlocking the value hidden in wastewater

Every day, wastewater treatment plants process enormous quantities of water containing valuable resources that are still largely underutilised. Beyond protecting water quality, modern facilities have the potential to recover nutrients, generate renewable energy and produce reclaimed water, transforming wastewater into a valuable resource.

According to Eurostat, approximately **68 billion m³** of municipal wastewater is treated across Europe every year. This wastewater contains an estimated **3 million tonnes of nitrogen, 0.5 million tonnes of phosphorus,** and **1 million tonnes of potassium**. This is enough to supply around **13% of the EU's fertiliser nutrient demand**.

68 billion m³

municipal wastewater cleaned in Europe per year

Contains enough nitrogen,
phosphorus and potassioum to supply **13%** of the EU's fertiliser nutrient demand.

The organic carbon in wastewater could generate approximately **9.5 billion m³ of methane annually**, equivalent to around **100 billion kWh of renewable energy**, enough to power roughly **25 million households**. At the same time, improved resource recovery could significantly reduce sewage sludge production from 15 to 11 million tonnes per year.

Realising this potential depends on operating treatment plants more intelligently, making better use of available data and supporting operators with advanced decision-making tools.

1.2 From wastewater treatment plants to water resource recovery facilities

Traditionally, wastewater treatment plants (WWTPs) have focused on one primary objective: removing pollutants to meet legal standards. Today's Water Resource Recovery Facilities (WRRFs) go a step further. They recover energy, nutrients and reusable water while continuing to meet stringent environmental standards. Rather than treating wastewater as a waste stream, they view it as a source of valuable resources that can be returned to the economy.

For plant operators, this shift means balancing multiple objectives simultaneously: maintaining treatment performance, reducing energy consumption, recovering resources, minimising greenhouse gas emissions and ensuring stable operation under continuously changing conditions. Managing these competing priorities exceeds the capabilities of conventional control strategies alone and requires new digital tools capable of supporting faster and more informed operational decisions.

1.3 Why smarter operation remains challenging

Although wastewater treatment plants generate large volumes of operational data, turning that data into reliable operational intelligence remains a significant challenge. Three key barriers continue to limit the wider adoption of advanced AI-based decision support and control.

Complex processes require advanced operational support

Wastewater treatment involves highly dynamic biochemical and hydraulic processes that continuously influence one another. Conventional mechanistic models accurately describe these processes but require specialist expertise and are often too computationally intensive for everyday operational use. Plant operators therefore need tools that combine scientific accuracy with practical usability.

Reliable data is often limited

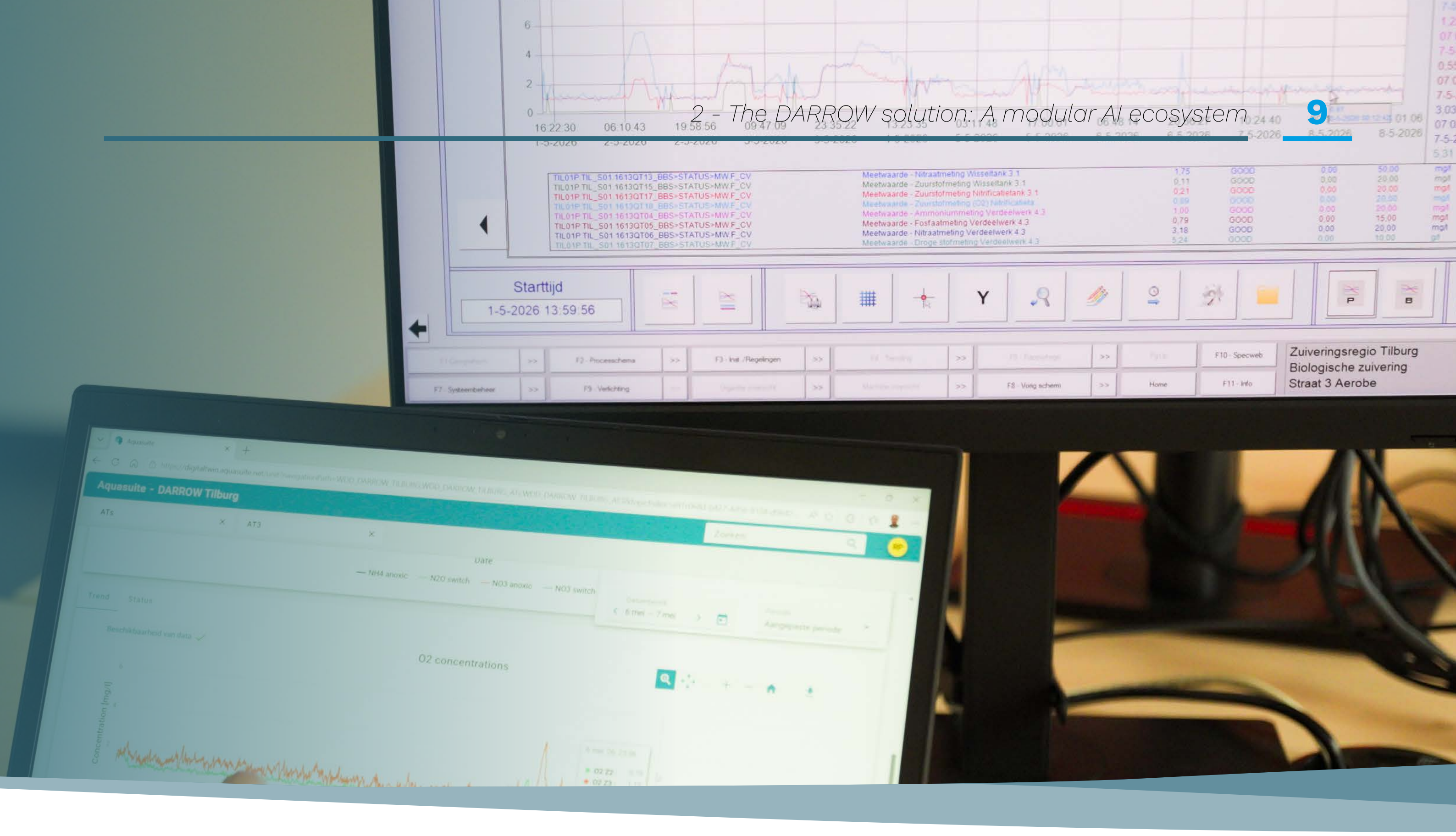
High-quality data is the foundation of any intelligent control system, yet wastewater treatment plants operate in harsh environments where sensors are exposed to fouling, clogging, drift and occasional failure. Many important process variables are not measured continuously, while laboratory analyses are typically performed only periodically because of their cost and complexity. As a result, data-driven models must be trained on incomplete, noisy and sometimes unreliable data.

AI must be transparent and trusted

While artificial intelligence offers significant opportunities for improving plant performance, operators must be able to understand and trust its recommendations before relying on them in daily operation. In critical infrastructure such as wastewater treatment, AI solutions must therefore provide transparent explanations, remain under human supervision and demonstrate robust performance under changing operating conditions.

Together, these challenges have limited the widespread adoption of advanced AI technologies in wastewater treatment plants. DARROW addresses each of them through an integrated approach that combines data enrichment, intelligent process control, digital twins and explainable AI to support operators in making better operational decisions.

> *The transition from wastewater treatment to resource recovery depends not only on better infrastructure, but on better use of data.*



02

The DARROW solution: A modular AI ecosystem

Digital technologies are transforming the way wastewater treatment plants operate, but most existing solutions address only individual challenges. Some focus on improving data quality, others on process optimisation or decision support. DARROW brings these capabilities together in a single, integrated AI ecosystem designed specifically for Water Resource Recovery Facilities.

Rather than replacing existing control systems, DARROW complements them by improving data quality, supporting operational decisions and enabling more intelligent process control. Its modular architecture allows utilities to adopt individual components independently or deploy the complete solution as their digital maturity evolves.

2.1 DARROW's two-layer architecture

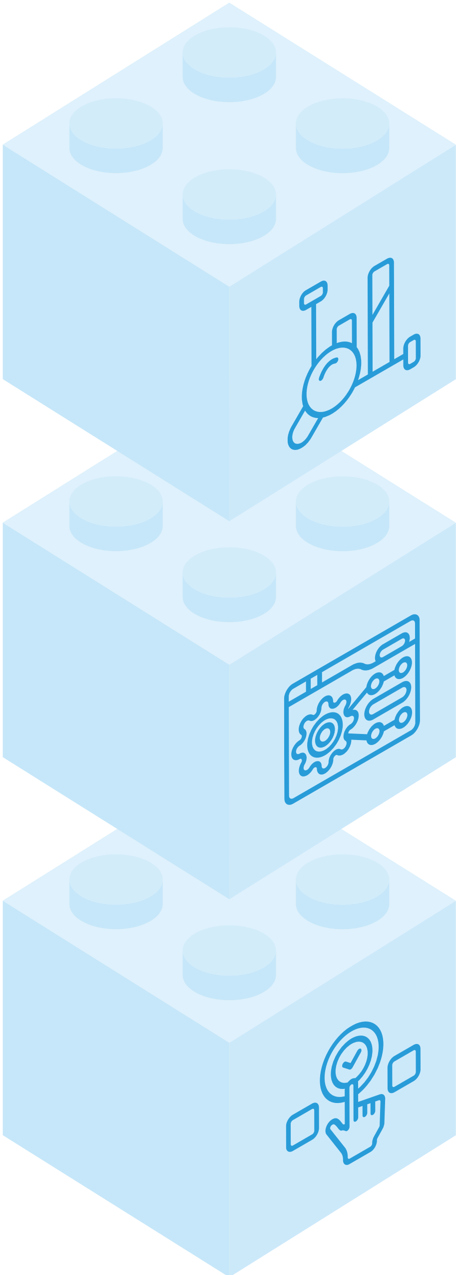
The DARROW solution is built on a two-layer AI architecture designed to bridge the gap between raw data and operational intelligence.

The first layer focuses on creating reliable information. It improves the quality, availability and reliability of plant data through data enrichment, software sensors and automated data validation.

Building on this trusted data foundation, the top layer delivers operational intelligence. It provides either operational recommendations via Digital Twins or executes autonomous control actions through advanced controllers.

2.2 Three complementary technology blocks

DARROW combines three complementary technological building blocks that work together to improve plant performance.



Smarter data

Machine learning techniques and Kalman Filters improve the availability and quality of operational data by generating missing information, detecting sensor anomalies and estimating variables that cannot be measured directly. This creates a reliable digital representation of plant operation and provides the foundation for all higher-level services.

Smarter control

Reinforcement Learning (RL) controllers provide safe and autonomous real-time optimisation for secondary treatment and anaerobic digestion. Unlike conventional controllers, they adapt to changing operating conditions and seek the best balance between treatment performance, energy consumption and resource recovery.

Smarter decisions

Digital Twins based on Reduced Order Models (ROMs) provide operators with fast, interactive decision-support tools. They predict future plant behaviour, evaluate alternative operating strategies and recommend optimal actions before changes are implemented in the real facility.

Together, these three capabilities allow operators to move from reacting to plant behaviour towards proactively managing treatment processes.

2.3 Design principles: modular, replicable, explainable

A key objective of DARROW was to develop technologies that could be deployed in real operational environments rather than remaining research prototypes.

The platform therefore follows three core design principles.

Modular
Each component can operate independently, allowing utilities to adopt only the capabilities they need. A plant may begin with data validation services and later expand to digital twins or autonomous control without redesigning the entire system.

Explainable
DARROW was designed with operators at its centre. Alongside technical development, the project investigated explainable AI approaches and worked closely with end users through workshops and evaluations to understand how AI can best support operational decision-making. These activities informed the design of trustworthy, human-centred AI solutions.

Replicable
Although developed and demonstrated at the Tilburg Water Resource Recovery Facility, the DARROW architecture has been designed for transfer to other treatment plants facing similar operational challenges. Its modular structure also makes the approach applicable to other complex industrial processes where reliable data, intelligent optimisation and human oversight are equally important.

03

Smarter data:
Augmentation and enrichment

High-quality data is the foundation of reliable AI and advanced process control. However, wastewater treatment plants often operate with incomplete laboratory measurements, noisy sensor signals and variables that cannot be measured continuously. DARROW addresses these challenges through a suite of data augmentation and enrichment services that improve data quality, fill information gaps and provide reliable estimates of key process variables.

3.1 Generating synthetic influent data

One of the main challenges in wastewater treatment is that key influent characteristics are not always measured continuously. At Tilburg WRRF, laboratory analyses provide measurements of parameters such as COD, TKN, TSS, BOD, NH_4, NO_3, TP and PO_4. However, laboratory analyses are available only at discrete sampling intervals, making it difficult to capture the full dynamics of the treatment process. The long intervals between measurements make it difficult to train and operate data-driven models.

To address this, DARROW developed predictive models that generate realistic time series and provide 24-hour forecasts of influent quality and quantity. Using Artificial Neural Networks trained on historical sensor data, rainfall information and operational data from Tilburg, the models predict inflow as well as COD, CODp, total nitrogen, total phosphorus, TSS, NH_4, NO_3 and BOD.

These predictions are available to operators through the DARROW platform and provide valuable input to ROM models. These predictions help improve the ROM models' performance and can help operators anticipate upcoming influent changes and improve operational planning.

3.2 Automatic sensor anomaly detection

Reliable sensor data is essential for intelligent plant operation. However, sensors in wastewater treatment plants are routinely affected by fouling, clogging, drift and other disturbances that reduce measurement quality.

At Tilburg, all sensor data arrives as raw I-Historian output without quality labels. Manual inspection is both time-consuming and impractical when analysing large datasets.

DARROW therefore developed an automatic data validation framework combining univariate and multivariate anomaly detection methods. While the univariate approach identifies abnormal behaviour in individual sensor signals, the multivariate approach evaluates relationships between multiple sensors simultaneously, providing a more comprehensive assessment of data quality.

Validated datasets generated by these methods form a reliable basis for downstream AI applications.

3.3 Real-time plant status classification

Understanding the biological state of a wastewater treatment plant is essential for stable and efficient operation, yet the microbial populations responsible for treatment cannot be measured directly in real time.

The DARROW Biomass Estimator addresses this challenge by predicting the concentrations of three key microbial groups (heterotrophic bacteria, autotrophic bacteria and polyphosphate-accumulating organisms) using a Long Short-Term Memory (LSTM) neural network trained on hundreds of simulated case studies using two weeks of historical sensor data for inference. This tool fills the gap left by the industry-wide absence of hardware-based biological sensors.

This level of visibility into the plant's internal status allows operators to be more precise and adaptive in their management strategies, moving beyond simple surface-level data. Ultimately, these insights into the microbial "engine" of the plant enable the optimisation of nutrient removal (nitrogen and phosphorus), leading to significantly improved treatment stability and efficiency while simultaneously minimizing energy consumption.

3.4 Software sensors for unmeasured variables

Many process variables that are important for plant operation cannot be measured continuously using conventional hardware sensors. To overcome this limitation, DARROW combines software sensors based on an Extended Kalman Filter (EKF) and machine learning algorithms.

The EKF operates through a continuous prediction-correction cycle, integrating real-time sensor streams with a simplified ASM2d mathematical model to provide reliable, real-time estimates of NH_4 and NO_3 concentrations. Crucially, the EKF goes beyond surface-level data to infer “hidden” or unmeasured state variables, such as maximum nitrification and denitrification capacities, which are vital for understanding biomass health and overall process stability.

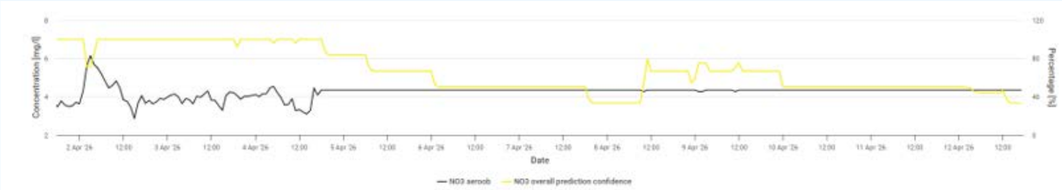
DARROW also incorporates specialized machine learning-based soft sensors to address one of the industry’s most significant environmental challenges: N_2O (nitrous oxide) emissions. As a potent greenhouse gas, N_2O is a critical Key Performance Indicator (KPI) for facilities aiming to meet Green Deal targets, yet direct hardware measurement is often prohibitively expensive and technically difficult. To overcome these barriers, the DARROW solution provides a real-time N_2O concentration estimate in both anoxic and aerobic zones and it also generates a comprehensive emission risk index.

Case study:

Automatic sensor anomaly detection at the Tilburg WWRF

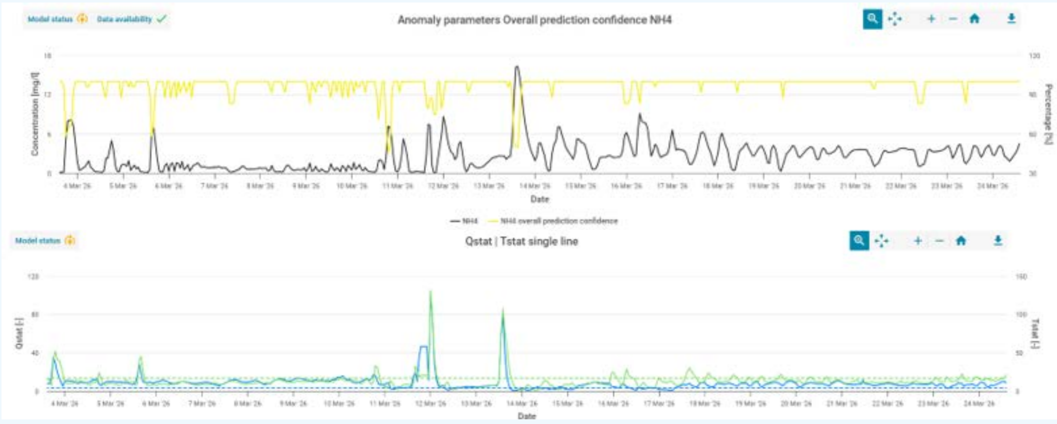
Example #1: Detecting a failing NO_3 sensor

A drop in the confidence value of the NO_3 measurements by the univariate detection model indicated a clear fault on the NO_3 sensor from April 4th onwards. This situation led to different operational settings that also affected the DO measurements.



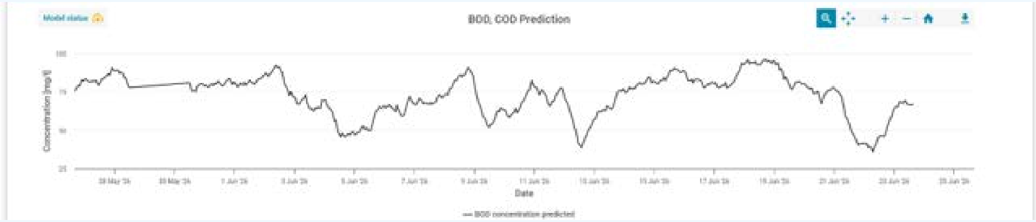
Example #2: Detecting abnormal NH_4 concentrations

Between March and April, the univariate anomaly detection model detected several periods in which the NH_4 concentration was higher than usual. Discussions with operators at De Dommel confirmed that these events were caused by an industrial pollution discharge entering the plant. The multivariate detection method independently identified the same period, providing additional confidence in the diagnosis.



Example #3: Real-time BOD predictions

DARROW's influent prediction models continuously forecast incoming BOD concentrations and display the results through the project platform. These forecasts help operators anticipate changing treatment loads and adjust aeration strategies in advance, improving both process performance and energy efficiency.



04

Smarter control: Reinforcement learning controllers

Conventional control systems have enabled wastewater treatment plants to operate reliably for decades. However, increasing operational complexity, stricter environmental requirements and the need to optimise multiple objectives simultaneously call for more adaptive control strategies.

DARROW addresses this challenge through Reinforcement Learning (RL) controllers that continuously learn how to optimise plant operation while balancing treatment performance, energy consumption and resource recovery.

4.1 Why conventional control is not enough

At Tilburg, key process variables are currently controlled through a set of independent conventional control loops. While robust and well established, this approach has several limitations.

First, the control loops operate independently and cannot account for the complex interactions between treatment units in a non-linear, high-dimensional system. Second, each loop relies on only one or two sensor signals, making control vulnerable to faulty measurements while leaving much of the available process information unused. Finally, conventional controllers are typically tuned for a specific operating condition and may gradually become sub-optimal as influent characteristics and biological conditions change.

DARROW's Reinforcement Learning (RL) controllers are designed to overcome these limitations by continuously considering the overall plant state and adapting control decisions to changing operating conditions.

4.2 How reinforcement learning controllers work

A RL controller learns a control strategy by repeated interaction with a simulator of the treatment plant. During training, it tests different control actions, observes the resulting plant state and gradually learns which actions achieve the best overall performance.

Unlike conventional controllers that optimise individual control loops, the RL controller considers all available sensor information simultaneously, recognises temporal patterns, makes plant-wide decisions and keeps adapting as conditions change. Deep RL makes this feasible for the high-dimensional, partially observed control problem a WRRF represents. To train robust strategies, the controllers are exposed to a wide range of operational scenarios (typical, atypical and extreme influent and load conditions) generated with the mechanistic model.

4.3 Secondary Treatment controller

The DARROW Secondary Treatment RL controller optimises two key operational variables simultaneously:

- Aeration flow rate in the nitrification zone by dynamically adjusting DO setpoints based on incoming load.
- Internal recirculation flow rate, controlling nitrate return to the anoxic zone.

The key advance over current practice: all variables are optimised simultaneously as a single multi-input multi-output (MIMO) decision, rather than through independent loops.

4.4 Anaerobic Digester controller

The key operational challenge for the Tilburg anaerobic digester is the instability caused by varying organic sludge loads. Varying feeding can destabilise the digester and reduce biogas production. The DARROW AD controller addresses this challenge by:

- controlling the feeding flow rate to maintain a constant, stable load on the digester;
- balancing primary sludge feed through the CAMBI thermal hydrolysis unit to compensate for varying loads;
- dynamically optimising methane production depending on the operational context.

4.5 From simulation to safe deployment

Deploying an untested RL controller directly on a critical-infrastructure plant is not acceptable. DARROW therefore trains and tests the controllers first in a calibrated mechanistic simulator of the Tilburg treatment plant. This simulator provides a safe environment in which the agent can explore freely without operational risk. Historical plant data is then used to close the gap between simulation and reality before any live operation.

During live operation, the controller only issues setpoints within the constraints allowed by the PLC/SCADA system.

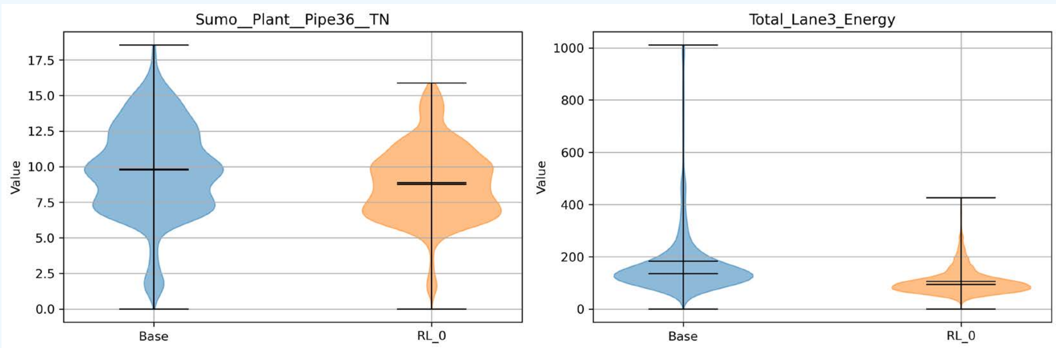
Case study:

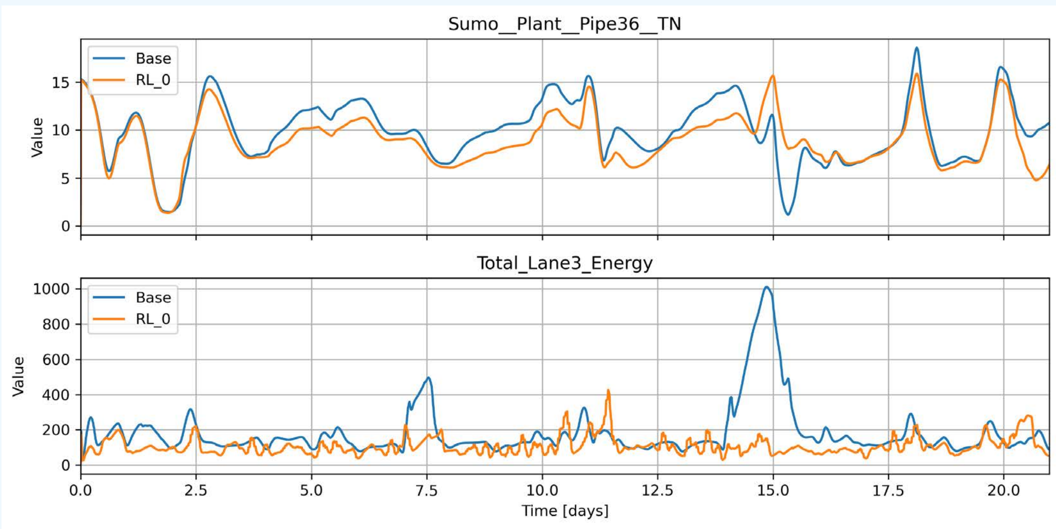
RL controller for Secondary Treatment in the WWRF Tilburg

Before DARROW, the biological treatment process at Tilburg WRRF relied on conventional control strategies. Each aeration zone was controlled independently by regulating dissolved oxygen (DO) setpoints, while the internal recirculation flow was adjusted based on the nitrate (NO_3^-) concentration in the anoxic tank. The combination of these setpoints depended largely on operator knowledge and experience to achieve effective nitrogen removal while minimising the energy required for aeration and pumping.

DARROW introduced a Reinforcement Learning (RL) controller that learns an optimal control policy from a calibrated process model. Instead of optimising aeration and recirculation separately, the RL controller coordinates both control actions simultaneously and recommends the combination of setpoints that best meets the operational objectives under changing plant conditions.

During the demonstration at Tilburg WRRF, the RL controller achieved X% lower energy consumption while maintaining/improving effluent nitrogen concentrations, demonstrating the potential of AI-assisted control to enhance process performance and energy efficiency.





05

Smarter decisions: Digital Twins for prediction and process optimisation

Reliable operational decisions depend on understanding how a treatment plant is likely to respond to changing conditions. However, testing different operating strategies directly on a live facility is neither practical nor risk-free.

DARROW addresses this challenge through a suite of Digital Twins that combine mechanistic knowledge with machine learning. These fast, interactive models allow operators to predict plant behaviour, evaluate "what-if" scenarios and identify optimal operating strategies before implementing changes in the real plant.

5.1 From complex mechanistic model to fast operator tool

Traditional mechanistic simulators are the gold standard for accuracy but can be too computationally demanding for running complex "what-if" scenarios in a real-time environment.

DARROW bridges this gap using a hybrid workflow: we use high-fidelity mechanistic simulators to generate thousands of "physics-consistent" scenarios, which then train fast machine-learning surrogates (ROMs). These ROMs retain the dominant dynamics of the complex model while enabling rapid simulation. This speed transforms a heavy engineering tool into a fast, interactive dashboard for daily operational decisions.

5.2 Secondary Treatment Digital Twin

The Secondary Treatment (ST) Digital Twin, implemented in a cloud-based and easily accessible architecture, serves as a powerful decision-support tool for facility operators. This platform provides three core functionalities: prediction, testing, and recommendation. It is designed to bridge the gap between complex modeling and daily operations.

Built using a Long Short-Term Memory (LSTM) neural network that recognizes complex temporal patterns, the tool accurately predicts nitrogen-related effluent indicators (NH_4, NO_3, TN). It empowers operators to test "what-if" scenarios by immediately simulating the impact of changing oxygen setpoints or recirculation rates on effluent quality. Furthermore, the system incorporates an optimisation algorithm that recommends the most efficient and feasible operating profiles for the next 24 hours, ensuring regulatory compliance while avoiding abrupt, disruptive control changes.

5.3 Anaerobic Digester Digital Twin

The Anaerobic Digester (AD) Digital Twin focuses on optimising the “sludge-to-energy” pathway. By analysing three years of historical data from the Tilburg WRRF, this XGBoost-based model predicts methane production at both hourly and daily timescales. These forecasts provide operators with day-ahead information, allowing them to optimise sludge feeding strategies to maximise biogas yield while ensuring the biological stability of the digester.

5.4 Plant-wide Digital Twin

Wastewater treatment is a balancing act between water quality and energy consumption. The Plant-Wide Digital Twin couples the water-line and sludge-line models to provide a facility-scale perspective. Using a multi-objective optimization algorithm (NSGA-II), the tool generates a Pareto front—a set of optimal operating plans that highlight the trade-offs between a Water Quality Index (WQI) and an Energy Operational Index (EOI). Operators can choose from “quality-first,” “cost-first,” or “balanced” strategies based on current plant priorities.

5.5 Explainable AI: Why every recommendation comes with a reason

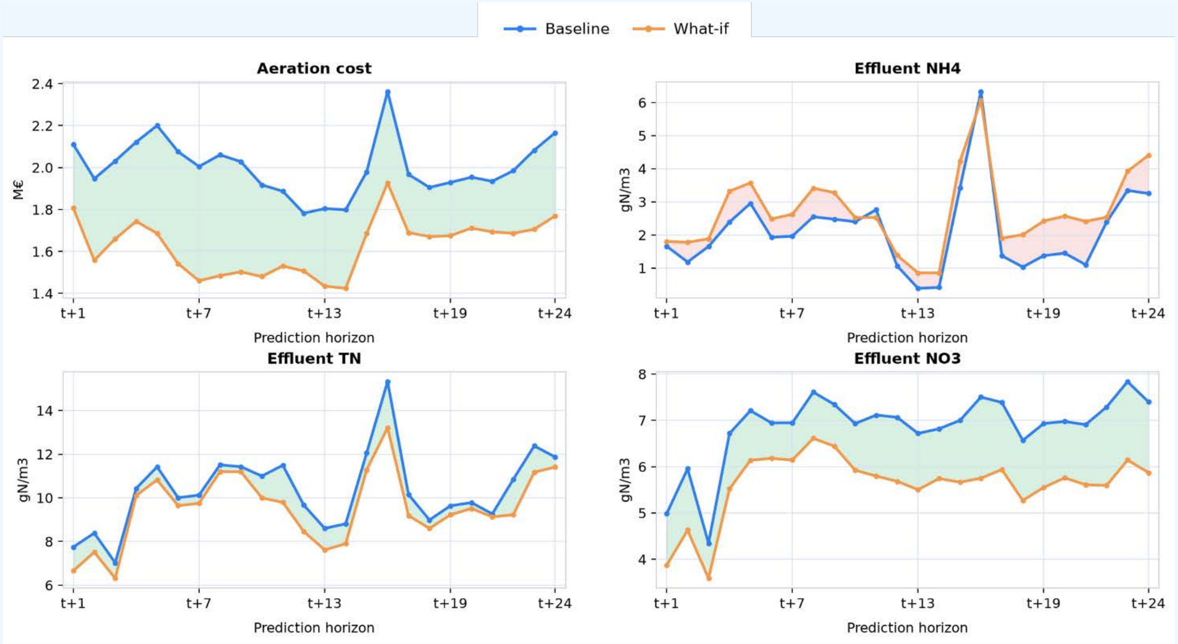
Operators will not act on recommendations they cannot understand or verify, and in a regulated, safety-critical setting they must be able to check the basis of any advice.

DARROW therefore incorporates an explainability layer to increase trust in the forecasts. Each prediction is accompanied by information showing which operational variables most strongly influenced the result, while a forward analysis (ICE/PDP) illustrates how predicted plant performance changes when individual operating parameters are adjusted.

Operator surveys carried out during the project confirmed a strong preference for transparent and understandable recommendations, with clear textual explanations slightly preferred over purely visual representations. These findings directly informed the design of the DARROW decision-support tools, ensuring that operators remain in control while benefiting from AI-assisted recommendations.

Case study:
Water-line ROM

If an operator is considering reducing energy costs, they can test a "what-if" scenario: *"If I switch off the facultative tank during this high-load period, what happens?"* The tool immediately predicts the specific increase in effluent ammonium and the corresponding decrease in aeration cost, allowing for a data-driven risk assessment.

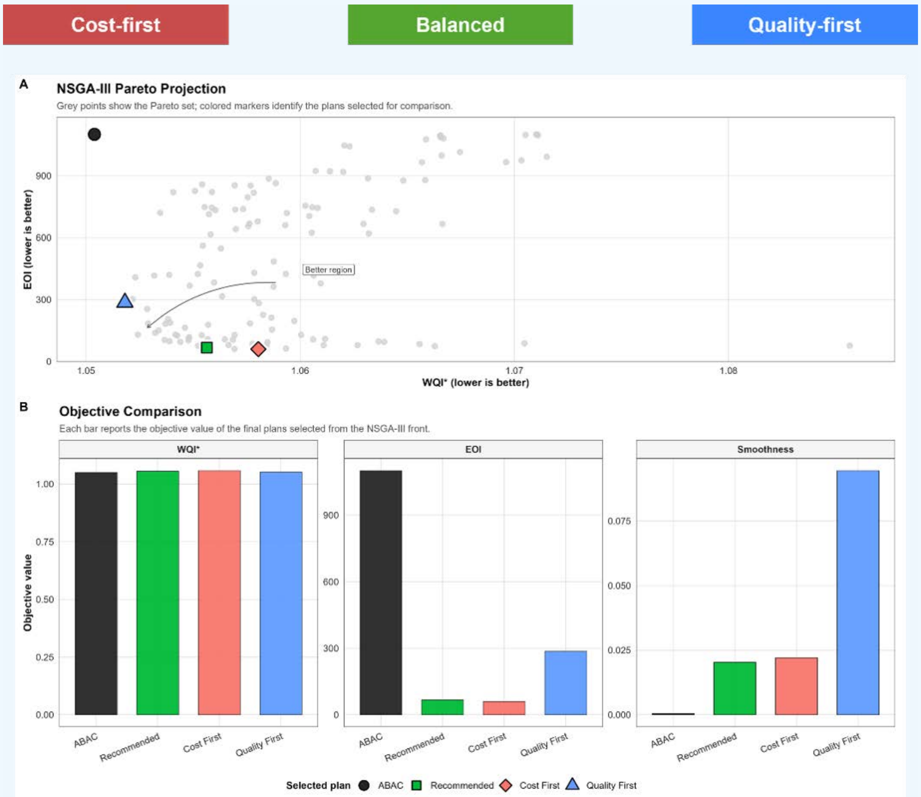


The baseline scenario considers dissolved oxygen setpoints of 2.0 gO_2/m^3 in the facultative tank and 2.0 gO_2/m^3 in both nitrification tanks. In the what-if scenario, the facultative tank aeration is switched off, while the nitrification tanks remain unchanged. The results indicate that switching off the facultative tank decreases aeration cost across most prediction horizons. This reduction also modifies the nitrogen dynamics: NH_4 tends to increase, while NO_3 decreases and TN remains similar or slightly lower. These results highlight the trade-off between energy savings, effluent ammonium and total nitrogen control.

Case study:

Plant-wide Digital Twin

Wastewater treatment is a balancing act between water quality and energy consumption. The Plant-Wide Digital Twin couples the water-line and sludge-line models to provide a facility-scale perspective. Using a multi-objective optimization algorithm (NSGA-III), the tool generates a three dimensional Pareto front, a set of optimal operating plans that highlight the trade-offs between a Water Quality Index (WQI) and an Energy Operational Index (EOI) and Smoothness for the aeration and recirculation setpoints. Operators can choose from "quality-first," "cost-first," or "balanced" strategies based on current plant priorities.



The optimisation algorithm identifies operational strategies that improve upon the baseline configuration by exploring trade-offs between effluent quality, energy performance and operational smoothness.

Panel A shows the Pareto front projection, where grey points represent feasible Pareto-optimal solutions and coloured markers indicate the selected plans for comparison. Panel B compares the objective values of the selected alternatives.

The Recommended strategy provides the best overall compromise, achieving low energy demand while maintaining effluent quality and moderate operational smoothness. In contrast, Cost First prioritises energy reduction, whereas Quality First favours process performance at the expense of higher operational effort.

DARROW

06

Demonstration at Tilburg: Results in practice

The DARROW technologies were developed and validated at the Tilburg WRRF, operated by Waterschap De Dommel in the Netherlands. The demonstration brought the DARROW tools into a real operational environment, allowing us to evaluate their performance using plant data and realistic operating conditions.

The demonstration also provided valuable feedback from operators, helping ensure that the developed tools address practical operational needs and can support future deployment at other treatment plants.

6.1 About the Tilburg demonstration site

Tilburg WRRF is the second largest wastewater treatment plant operated by Waterschap De Dommel, with a design capacity of 350,000 population-equivalent. Three parallel treatment trains handle wastewater flows of approximately 2,000 m^3/h during dry weather and up to 12,585 m^3/h during wet weather. The plant includes primary clarifiers, biological treatment, chemical phosphorus removal, and secondary clarifiers in the water line. The sludge line includes a CAMBI thermal hydrolysis unit, anaerobic digesters, a Phospaq struvite recovery unit and a side-stream treatment unit, all collectively referred to as the Energy Factory Tilburg.

The primary operational objective is reliable effluent quality at reasonable cost: total nitrogen below 10 mg/L and total phosphorus below 1 mg/L, while minimising energy consumption, chemical use, and greenhouse gas emissions.

During the DARROW project, the Tilburg WRRF undewernt three major infrastructure changes, including a new biogas-to-natural-gas upgrading unit (adding 20% more sludge capacity), an aeration tank expansion, and a tag code restructure. These changes were incorporated into the project's data management plan.

6.2 What was deployed and tested

DARROW validated technologies across all three functional areas of the platform (smarter data, smarter control, smarter decisions): the data-augmentation services (synthetic lab data, anomaly detection, plant-status classification and software sensors) on the live data streams; the RL controllers for secondary treatment trained on the calibrated simulator; and the ROM-based Digital Twins for prediction and recommendation.

6.3 Results: DARROW impact at Tilburg

The Tilburg demonstration confirmed the technical feasibility of deploying DARROW's AI-based services in a real wastewater treatment plant, with their impact at the Tilburg WRRF assessed using the key indicators summarised below.

Indicator	What is assessed
Nitrogen removal efficiency	The effect of coordinated aeration and internal recirculation control on nitrogen removal and effluent quality
Sludge production	The influence of plant operation on total sludge production and handling requirements
Energy consumption	Energy use for aeration, pumping and other major treatment processes
Energy generation	Renewable-energy production from biogas recovery
Greenhouse-gas emissions	Estimated nitrous oxide emissions from the biological treatment process

For the latest quantified results, visit the DARROW website.

07

Social impact

The successful adoption of artificial intelligence depends not only on technological innovation, but also on the people who use it. Throughout DARROW, particular attention was given to equipping operators with the knowledge needed to work confidently with AI-based tools and to understanding how such technologies can best support, rather than replace, human decision-making.

7.1 Training operators to work alongside AI

To support the adoption of AI in the water sector, DARROW developed a free, self-paced Massive Open Online Course (MOOC) through UGENT's CAPTURE Academy.

The course is designed for wastewater professionals, plant operators, non-expert users of AI and data systems, as well as Master's and PhD students. Delivered in English through the Thinkific platform, it consists of five modules covering:

- data quality and standards;
- hybrid modelling and Reduced Order Models;
- wastewater treatment control, including Reinforcement Learning;
- optimisation using Digital Twins and ROMs;
- the DARROW platform, tools and practical case studies.

Each module combines video lectures, reading material and interactive exercises, and completion is recognised with a digital certificate.

7.2 Building operator trust in AI

The introduction of AI into critical infrastructure depends on operator confidence and acceptance. DARROW therefore explored how wastewater professionals perceive different levels of automation and which factors influence their willingness to adopt AI-supported decision-making.

The assessment compared three levels of automation: systems that provide warnings, systems that make recommendations and systems capable of acting autonomously.

Operators rated all three as moderately useful, but the most autonomous option drew clearly less confidence, enjoyment and willingness to adopt. There is a strong preference for keeping the human operator in control.

Across all scenarios, operators expressed a strong, consistent desire for transparency and sufficient explanation of how recommendations are generated. Clear textual explanations were slightly preferred over purely visual formats. Operators also asked for features such as expected-outcome information, a way to give feedback when they disagree, and clarity on override authority. These early findings (from a small operator sample) directly shaped DARROW's explainability layer and human-in-the-loop design.

> *"AI should support operators, not replace them."*

08

Lessons learnt

Developing and deploying AI technologies in an operational wastewater treatment plant provided valuable technical and organisational insights. Beyond validating the DARROW technologies, the project highlighted several practical considerations that will help guide future implementations in the water sector.

8.1 Choosing the right reference data for performance evaluation

Evaluating the impact of advanced control strategies is not always straightforward. Comparing current plant performance with historical operation may be misleading because wastewater characteristics, weather conditions and plant configuration can change significantly over time.

Within DARROW, performance was therefore assessed by comparing parallel treatment lines operating under the same conditions. This approach made it possible to isolate the effect of the new control strategy while minimising the influence of external factors.

8.2 Advanced controllers require appropriate modelling approaches

When advanced controllers (such as an RL agent) are applied, modelling and simulation with a mechanistic model become more difficult. Additionally, using a mechanistic model as the basis for a learned controller raises real tuning and acceptance-gate challenges. It may be feasible, but it is demanding and not always as fit-for-purpose or economical as conventional approaches. This is a trade-off that treatment plants carefully weigh.

8.3 Cybersecurity must be considered from the beginning

As wastewater treatment plants form part of Europe's critical infrastructure, advanced digital solutions must comply with strict cybersecurity requirements, including those introduced under the NIS2 Directive.

DARROW therefore incorporates several safeguards within its

functional architecture: the system only issues advisory setpoints within PLC/SCADA-allowed constraints; a validity/quality status accompanies every calculated value; watchdog functionality warns the operator (or switches from advanced control back to the PLC) when values cannot be trusted; all inputs, outputs and setting changes are logged for traceability; and information-security management follows ISO/IEC 27001.

Together, these measures keep humans in oversight and make the system auditable by authorities.

8.4 Start with a minimal viable product

Focusing on a working prototype with a real technical implementation, rather than an exhaustive feature set, helped the DARROW project deliver demonstrable value at the Tilburg plant and learn from real operation early on.

8.5 Good AI starts with good data

Advanced AI solutions depend on reliable operational data. During the project, it became clear that the level of instrumentation, sensor quality and availability of historical data strongly influence both model development and deployment.

This highlights an important message for future adopters: investing in reliable data infrastructure is often the first step towards successful digitalisation.

09

Could DARROW work at your plant?

DARROW has been designed as a modular solution that can be deployed progressively according to the needs and digital maturity of individual wastewater treatment plants. While the complete platform offers the greatest benefits, many of its individual services can also be implemented independently, allowing utilities to build capabilities step by step.

9.1 Technical prerequisites

Successful deployment of DARROW depends on access to reliable operational data and a minimum level of digital infrastructure. Table 2 summarises the instrumentation and infrastructure requirements for deploying DARROW.

Requirement	Requirement level	Notes
Online DO sensors in aeration tanks	Essential	Required for RL controller
Online NH_4 sensor (influent or effluent)	Essential	Required for software sensors
Online NO_3 sensor	Essential	Required for RL
Flow meters (influent, recirculation, sludge)	Essential	Required for all models
Temperature and pH sensors	Essential	Minimum baseline instrumentation
Phosphate online sensor	Recommended	Improves P removal recommendations
SCADA / PLC system with data historian	Essential	For data access and actuation
Reliable internet / cloud connectivity	Essential	Required for top-layer services
COD online sensor (influent)	Optional	Improves influent load forecasting
N_2O online sensor	Optional	Required for N_2O risk service (CWEU)

9.2 The deployment pathway

DARROW deployment follows five phases, typically over 12-18 months:

1. Data-access audit and API set-up, connecting to the plant's historian and laboratory systems.
2. Mechanistic-model calibration on historical plant data.
3. Bottom-layer services deployment and validation (data augmentation, anomaly detection, software sensors).
4. RL-controller and ROM-based digital twins deployment and testing on a virtual emulator of the plant.
5. Live deployment with operator training and human oversight.

9.3 Start small and scale up

Plants do not need to deploy the full DARROW platform from the outset.

The data enrichment services provide immediate value with relatively limited integration effort and can serve as a practical entry point. As confidence grows and operational requirements evolve, additional components such as Digital Twins and Reinforcement Learning controllers can be introduced.

This modular approach reduces implementation risk, allows organisations to build experience gradually and supports long-term digital transformation.

10

Scaling up: Replication across Europe and beyond

The successful demonstration of DARROW at Tilburg provides a foundation for wider deployment across Europe. The project's modular architecture, combined with its focus on practical implementation, positions DARROW as a scalable solution for wastewater treatment plants seeking to improve operational efficiency, resource recovery and sustainability.

10.1 Market opportunity

The first commercialisation phase targets WRRFs across the EU and UK. These facilities serve around 500 million people. The DARROW platform suits only large WRRFs (>100,000 PE) that also have the necessary process-automation and sensors. This reduces the serviceable addressable market to about 400 plants, treating wastewater of roughly 150 million people, with an aimed market share of 25%. Front-runner utilities already identified as potential early adopters are Waterschap Vallei & Veluwe (NL), Northern Ireland Water (UK) and Aquafin (BE), alongside Waterschap De Dommel.

The consortium is already building on the knowledge generated within DARROW to deploy Digital Twin technologies at full scale, including implementations led by Ceit in Spain. These early deployments demonstrate the project's potential for wider replication and commercial uptake.

The DARROW business model combines direct sales (onboarding plants at low margin to lower the entry threshold) with subscription/licensing; the integrated AI toolbox is marketed through RHDHV, with developer partners compensated via licences/royalties and also free to exploit their own components.

10.2 Expected scale-up impact by 2030

If deployed at 60 large wastewater treatment plants within five years of project completion, DARROW is expected to deliver significant environmental and economic benefits. Table 3 summarises the projected impact.

Outcome (5 years after project end)	Target value
WWTPs implementing the DARROW solution	60 (>100,000 PE each)
Water treated by those plants (Hm^3/year)	1,000
Increase in biological phosphorus recovered (t P/year)	4,000
Reduction in energy consumption (MWh/year)	92,000
Increase in energy generation (MWh/year)	27,600
Reduction in GHG emissions (t CO_2eq/year)	25,000
Reduction in sludge waste production (m^3/year)	238,900
Forecast revenues from exploitation (€)	15 million
Direct jobs created	25

10.3 Beyond the water sector

Although DARROW was developed for wastewater treatment, its underlying architecture is applicable wherever complex industrial processes combine large volumes of operational data with dynamic physical and biological systems.

The combination of mechanistic modelling, Reinforcement Learning, Digital Twins and Explainable AI can support optimisation in sectors such as food and beverage production, biogas and bio-refineries, and the chemical process industry.

This flexibility extends the potential impact of DARROW beyond the water sector and opens opportunities for future research, innovation and commercial deployment.

Scan to access more information on **www.wastewater.ai**

Project Coordinator

Dr. Beñat Elduayen Echave
Associación Centro Tecnológico Ceit
belduayen@ceit.es

Media, Press and Communication

Dr. Mihaela Bozukova
European Science Communication Institute gGmbH (ESCI)
mbo@esci.eu

Co-funded by the European Union