ToDrinQ

Final Five modules for improved water systems monitoring and operational support

Deliverable 7.2

WP7 Integrated modular platform for Drinking Water Systems’ monitoring and operational support

Funded by the European Union

GRANT AGREEMENT NUMBER	101082035		
FULL TITLE / ACRONYM	ToDrinQ		
START DATE	01-12-2022	DURATION	48 months
END DATE	30-11-2026		
PROJECT URL	www.todrinq.eu		
WORK PACKAGE No and title	WP7 Integrated modular platform for Drinking Water Systems' monitoring and operational support		
DELIVERABLE TITLE	Deliverable D7.2. Final five modules for improved water systems monitoring and operational support		
ACTUAL DATE OF DELIVERY	11-12-2025 (1) & 21-04-2026 (2) & 10-06-2026 (3)		
NATURE	R	DISSEMINATION LEVEL	Sensitive
LEAD BENEFICIARY	Delft University of Technology		
RESPONSIBLE AUTHOR	Luuk Rietveld, Delft University of Technology		
CONTRIBUTIONS FROM	Panagiotis Kossieris, NTUA Vasiliki Thomopoulou, NTUA George Bariamis, NTUA Greg Kyritsakas, Delft University of Technology Akrivi Alexandraki, Delft University of Technology Siddharth Seshan, KWR Jasper Immink, KWR		

Document history

Version	Description (Section, page number)	Author	Organisation short name
V0.1	First draft	All authors	Delft University of Technology
V0.2	Review	Alex van der Helm	Stichting Waternet
V0.3	Draft final 1 Processed of feedback received from reviewers / partners	All authors	Delft University of Technology
V1.0	Final version	Luuk Rietveld	Delft University of Technology
V2.0 & V3.0	Final version	Luuk Rietveld	Delft University of Technology

Quality control

Author	Organisation short name	Role	Date
Luuk Rietveld	Delft University of Technology	Deliverable Leader & Work Package Leader	11-12-2025 & 21-04-2026 & 10-06-2026
Alex van der Helm	Stichting Waternet	Reviewer 1	19-11-2025
Luuk Rietveld	Delft University of Technology	Scientific project coordinator	11-12-2025
Danitsja van Heusden	Delft University of Technology	Project coordinator	11-12-2025 & 21-04-2026 10-06-2026

Table of contents

ABBREVIATIONS 8

EXECUTIVE SUMMARY 9

1. INTRODUCTION 10

1.1 The ToDrinQ project 10
1.2 Objectives for Integrated modular platform for Drinking Water Systems' monitoring and operational support 10

2. IDENTIFICATION OF USER NEEDS AND SPECIFICATIONS FOR WATER SUPPLY OPERATORS 11

3. INDIVIDUAL MODULES FOR MONITORING AND OPERATIONAL SUPPORT 13

3.1 Introduction 13
3.2 Risk-based Source to tap water supply management 15
3.3 Catchment/source 22

3.3.1 Nutrient runoff 24

3.3.2 Concentrations of Chl-a, pH and DO 26

3.3.3 Turbidity at the source 28

3.4 Conveyance network 28

3.4.1 Turbidity risk assessment 31

3.4.2 Conductivity risk assessment 33

3.5 Performance and state of water treatment 34
3.6 Distribution network 37
3.7 Sensors in the Modules 40
3.8 Modules architecture (backend) DC#1 45

3.8.1 Backend architecture 46

3.8.2 Modules from Source to Tap 47

3.8.3 Integration of the soft sensors in the modules 50

3.8.4 Full pipeline of all models 50

3.8.5 Demonstration setup and dummy data implementation 64

3.8.6 Live prediction and user interface support 64

3.8.7 Transition to real deployment at WTNT 65

3.8.8 FIWARE alignment, NESSIE integration, and AWS deployment 65

3.8.9 Running the platform locally 66

4. CONCLUSIONS 68

4.1 Identification of user needs 68
4.2 Contribution to monitoring and operational support 68

4.3 Innovation and pathways to impact 69

List of figures

Figure 1: Format as basis for co-creation for specific production location 11

Figure 2: The integrated modular platform data and workflow.. 13

Figure 3: Water supply modules' representation expressed as Data Model 14

Figure 4: Looking inside the modules 15

Figure 5: Input and Output of the Risk Assessment Module. An example using synthetic data for one of the soft sensors in the Water Treatment Module is given 19

Figure 6: Example of decision tree for intervention pathways... 21

Figure 7: Key catchment elements of catchments characteristics to be considered 22

Figure 8: Mapping of catchment elements, potential hazards and potential risks 23

Figure 9: Time series of exceeding pixels and corresponding thresholds 28

Figure 10: Key catchment elements of conveyance systems characteristics to be considered 29

Figure 11: Mapping of conveyance elements, potential hazards and potential risks 30

Figure 12: Risk bands as indicated by the data analysis and operational turbidity levels........................ 32

Figure 13: Turbidity risk bands projected in the available time series 32

Figure 14: Risk bands as indicated by the data analysis and operational turbidity levels........................ 33

Figure 15: Conductivity risk bands projected in the available time series 34

Figure 16: Mapping of water treatment elements, potential hazards and potential risks 35

Figure 17: Mapping Chlorinated Distribution systems, potential hazards and potential risks 38

Figure 18: Mapping Non-Chlorinated Distribution systems, potential hazards and potential risks 39

Figure 19: Backend architecture of the ToDrinQ platform, showing data flow from SCADA/PIMS through the connector to the Dockerized FastAPI platform, where Source, Treatment, Distribution, and Risk modules are implemented.. 47

Figure 20: FastAPI Swagger interface showing the implemented endpoints of the **Source module**........ 48

Figure 21: FastAPI Swagger interface showing the **Treatment module** endpoints................................. 49

Figure 22: FastAPI Swagger interface presenting the **Distribution module** endpoint 49

Figure 23: FastAPI Swagger interface showing the **Risk module**, which aggregates outputs from all other modules.. 49

Figure 24: Full pipeline of the models on the platform and their connection is given (Part 1)................ 51

Figure 25: Full pipeline of the models on the platform and their connection is given (Continuation from Part 1 – Part 2) 52

Figure 26: FastAPI Swagger interface showing the Source module endpoints, including the source overview and the latest output of the turbidity inlet soft sensor (SoSe #9) 53

Figure 27: Example execution of the Source module endpoint in the FastAPI backend, showing the retrieval of the latest output from the turbidity inlet soft sensor (SoSe #9).......... 54

Figure 28 Treatment module which includes the Soft Sensors SoSe #8, 10, 10.1, 10.2, RL #5 and Hard sensor: BactoSense #4 55

Figure 29: FastAPI Swagger interface showing the Treatment module endpoints, including the coagulation prediction service.......... 55

Figure 30: Example execution of the coagulation prediction endpoint within the Treatment module, showing input parameters retrieved from the data source layer and the corresponding structured JSON output 56

Figure 31: Example execution of the ozonation CT prediction endpoint within the Treatment module . 57

Figure 32: Example execution of the Distribution module endpoint, showing the retrieval of the latest output from a microbiological soft sensor (SoSe #11) 58

Figure 33: Click on the risk module.......... 59

Figure 34: Choose for example the SoSe #9 which predicts the turbidity at the inlet. 60

Figure 35: Example response from the Risk module, illustrating the transformation of model outputs into structured risk information 60

Figure 36: Continuation of the response body. Part 2 of the Figure 35 61

Figure 37: Short and long term actions (output from the risk assessment module.......... 62

Figure 38: Decision tree - output of the risk assessment module 63

Figure 39: It presents the graphical user interface of the platform, where both manual and live prediction modes are available.......... 64

Figure 40: Deployment and integration architecture of the ToDrinQ platform, showing the connection with WTNT through PIMS and a dedicated connector, as well as the cloud-based deployment on AWS enabling interoperability with the NESSIE platform 66

List of tables

Table 1: List of ideas from the DC#1: Amsterdam, the Netherlands (WTNT), within the WP7 co-creation workshop 11

Table 2: Taxonomy of key elements of risk-based assessment containing physical components of the system, hazards and other risk factors as well as risk management and mitigation actors (regulation and stakeholders.......... 17

Table 3: Probability is translated to likelihood. There are three possible classes: low, moderate and high. 18

Table 4: Basis for decision tree analysis for risk-based interventions pathways.......... 20

Table 5: Hazard, risks and potential mitigation strategies for catchment quantity and quality issues.....24

Table 6: Discharge Classification.....25

Table 7: Monthly Nutrient Load Classification.....25

Table 8: Combined Risk Assessment Matrix.....26

Table 9: Critical thresholds for Chl-a, pH and DO.....26

Table 10: Risk Classification Thresholds.....27

Table 11: Hazard, risks and potential mitigation strategies for conveyance quantity and quality issues.31

Table 12: Turbidity risk bands.....31

Table 13: Conductivity risk bands.....33

Table 14: Hazard, risks and potential mitigation strategies for treatment issues.......**Error! Bookmark not defined.**

Table 15: Risk Identification for Chlorinated Distribution systems.....38

Table 16: Risk Identification for Non-Chlorinated Distribution systems......39

Table 17: Sensors and RL models on the modules of the platform DC #1 WTNT and DC #2 EYDAP.........41

Table 18 Mapping of #DC1 (WTNT) Sensors and Models to Platform Modules and Their Contribution to the Risk Assessment Module......43

Table 19 Mapping of #DC2 (EYDAP) Sensors and Models to Platform Modules and Their Contribution to the Risk Assessment Module......44

Abbreviations

AOC	Assimilable organic carbon
API	Application Programming Interface
AWS	Amazon Web Services
BrO_3^-	Bromate
DBP	Disinfection by-product
DC	Demo Case
DWT	Drinking water treatment
EU	European Union
$FeCl_3$	Ferric chloride
PFAS	Per- and polyfluoroalkyl substances
RL	Reinforced Learning
TRL	Technology Readiness Level
WQI	Water Quality Index
WP	Work Package
WSP	Water Safety Plan

Short name	Legal Name
TUD	TECHNISCHE UNIVERSITEIT DELFT
RWTH	RHEINISCH-WESTFAELISCHE TECHNISCHE HOCHSCHULE AACHEN
CEB	CENTRE BELGE D'ETUDE ET DE DOCUMENTATION DE L'EAU
NTUA	ETHNICON METSOVION POLYTECHNION
KWR	KWR WATER BV
WTNT	STICHTING WATERNET
HWL	HET WATERLABORATORIUM NV
VEF	VEOLIA EAU - COMPAGNIE GENERALE DES EAUX SOCIETE EN COMMANDITE PAR ACTIONS
VEOCZ	VEOLIA CESKA REPUBLIKA, A.S.
EYDAP	ETAIREIA YDREYSEOS KAI APOCHETEFSEOS PROTEYOYSIS ANONIMI ETAIREIA
OLI	OLISENS TECH
OXY	OXYMEM LIMITED
ORV	ORVION B.V.
CHEM	CHIMIKI TECHNOLOGIA P. DIMOPOULOU -P.TAZES & SIA OE
WE	WATER EUROPE
ALTIS	ALTIS Groupe SA
BNV	BNOVATE TECHNOLOGIES SA
VEHO	Veolia Holding Ceska Republika AS

Executive summary

The present Deliverable (D7.2) is part of Work Package (WP) 7 of the ToDrinQ project, titled *"Integrated modular platform for Drinking Water Systems' monitoring and operational support."* The objective of WP7 is to develop an integrated, modular, and flexible platform to support drinking water operators in monitoring and controlling their systems. WP7 consists of four tasks: (T7.1) identification of user needs, (T7.2) interoperability standards, (T7.3) development of individual modules, and (T7.4) development of the online platform.

While D7.1 presented the conceptual design, D7.2 delivers the implemented backend software and its operational architecture. The platform is realized as a FIWARE-aligned system supporting two demonstration cases: WTNT (DC#1) and EYDAP (DC#2). For WTNT, a Dockerized FastAPI backend (developed by TU Delft) is deployed within the existing infrastructure and connected to SCADA/PIMS systems for real-time data processing. For EYDAP, the NESSIE platform (developed by NTUA) is used. The two cases are integrated at project level through NESSIE, which communicates with the ToDrinQ backend via Application Programming Interface (API) requests following NGSI-LD standards.

C The modules are structured across the drinking water supply chain, including Source (Section 3.3), Conveyance (Section 3.4), Treatment (Section 3.5), and Distribution (Section 3.6), supported by a source-to-tap risk management framework (Section 3.2). Within each module, sensor outputs are translated into risk indicators categorized as low, moderate, or high based on the likelihood of threshold exceedance. These risk levels are directly linked to predefined mitigation actions, enabling timely operational decision-making.

For deployment, the WTNT backend is delivered as a Docker image and connected to real operational data via PIMS. In parallel, the same backend is deployed on Amazon Web Services (AWS) and integrated with NESSIE, enabling a unified environment where both demonstration cases are represented, while maintaining data confidentiality through the use of dummy data where required.

Overall, D7.2 demonstrates the transition from conceptual design to functional and deployable modules, supporting real-time monitoring, predictive modelling, and risk-based decision-making in line with the objectives of the ToDrinQ project and the requirements of the revised EU Drinking Water Directive.

1. Introduction

1.1 The ToDrinQ project

The ToDrinQ project is an EU funded project with an overall objective to support the implementation of the revised Drinking Water Directive and to significantly enhance the scientific and technical knowledge on drinking water quality protection, monitoring and treatment by developing and testing a compendium of modular complementary innovations (hereafter termed the 'ToDrinQ Toolkit'), including novel real time sensing and water quality parameters technologies, innovative treatment processes (especially suitable for modular, adaptable treatment plants) and interoperable, easy to deploy decision tools that support resilient, evidence-based treatment plant design and improved overall water system operational awareness and risk-based response. The ToDrinQ Toolkit aspires to increase the resilience of drinking water systems in terms of both increased robustness (against short term stresses) and adaptability (against longer term uncertainties) and ensure high-quality drinking water, minimising micro-pollutants, pathogenic micro-organisms and disinfection by-products (DBPs).

1.2 Objectives for Integrated modular platform for Drinking Water Systems' monitoring and operational support

The present Deliverable **Five (5) modules for improved water systems monitoring and operational support** is part of Work Package (WP) 7 of the ToDrinQ project, named **Integrated modular platform for Drinking Water Systems' monitoring and operational support**.

The overall aim of this WP is to develop an integrated (but modular and flexible) platform able to support drinking water systems operators in monitoring and controlling their systems. Specific objectives are:

- To identify user needs and specifications for monitoring and operational support from water supply system operators
- To design and develop innovative modules supporting operators, by providing alarms, system information and intervention pathways for risk mitigation
- To develop an integrative platform that will operationalise these modules and be deployed in pilots to test and validate.

WP7 consists of four Tasks:

1. T7.1 Identification of user needs and specifications for water supply operators
2. T7.2 Interoperability Standards
3. T7.3 Developing the individual modules
4. T7.4 Developing the online platform for module deployment

In the present Deliverable (D7.2) the first and third tasks are addressed. The elaboration of T7.2 will be and T7.4 will further be described in D7.3 and D7.4.

2. Identification of user needs and specifications for water supply operators

During the first task (T7.1) an inventory was made, under the coordination of TUD, in relation to the performance of existing systems and available tools and methods. End-users (WTNT, EYDAP) were asked to fill out the form presented in Figure 1. Further, half a day workshops were organised with the end-uses to identify user needs and specifications for water supply operators, an essential step towards co-creation for future development. Requirements were then set-up for the development of tools to better integrate data, transform them to information and use the information for identification of hazards and related risks, predictions and possible interventions.

1. What is the goal of the production location
2. Description of the treatment facilities and related manipulations
3. Disturbances in
    - Production flow
    - Water quality
    - operation
4. Online measurements (hard sensors) for daily operation
5. Online soft sensing for daily operation
6. Laboratory measurements (for daily operation)
7. Process inspection
8. Use of information for operation
9. Use of DSSs/software for operation (in water supply)
10. Additional observations and possibilities for improvement (e.g. operation irt objectives; use of data for operation; missing data/information; measurements for state estimation; visibility of information for operation; etc.)

Figure 1: Format as basis for co-creation for specific production location

Table 1: List of ideas from the DC#1: Amsterdam, the Netherlands (WTNT), within the WP7 co-creation workshop

Idea No	Description	Score
5	Estimation of state of activated carbon filters for better use of total absorption capacity, e.g. for PFAS removal by using more/other filters in first/second stage (Leiduin DWTP)	100
3	Determination of dynamic setpoint for improved performance of coagulation (Dunea, PWN, WTNT and Evides/DPWE soft sensor for coagulation)	82
24	High resolution water demand prediction (at district/household level) to minimize necessity of smart meters for raising awareness and promote customer water saving	69
24	Scenarios, monitoring and (and consequences for) control DWT (e.g. on ozonation) when drinking water source (partly) changes, e.g. instead of using dune water, partly using directly pre-treated surface water. With purpose to reduce PFAS load (higher from Dune water that raw water) and to optimize max infiltration permits.	68
27	Stimulate alertness in operation by asking every day a technological question (generated by A.I.) to staff	61

Table 1 presents the five highest-scoring ideas, representing the main outcomes of the stakeholder discussions at WTNT. During the workshop, stakeholders further evaluated and refined these ideas, ultimately identifying that the key dynamics relevant for ToDrinQ project are centered around the coagulation treatment process and monitoring and control of Drinking Water Treatment (DWT). As a result, soft sensors and reinforcement learning models have been developed focusing on coagulation and ozonation. Overall, this contribution supports proactive decision-making within the ToDrinQ project. By enabling accurate estimation of coagulant and ozone requirements, the system enhances drinking water quality and helps reduce the formation of disinfection by-products in subsequent treatment stages.

Based on the stakeholder engagement at EYDAP, specifically for the integrated modular platform (WP7, based on NESSIE), the following topics were identified of being of interest:

- Polydendri DWTP (EYDAP DC#2) will treat water from Lake Yliki and from a borehole of Mavrosouvalas pumping stations in the near future (in 2024-2025) for a short period due to retrofitting works in the Mornos aqueduct. However, the borehole water is high in hardness and turbidity and will be used to mix groundwater and surface water of Yliki lake.
- Prediction of the water consumption is necessary as on some occasions a 10% increase in water consumption could result in extra treatment requirements, such as an extra tank of aluminium sulfate.
- How can we control the dosages of the chemicals dosed to the water for disinfection, automatically and not based on the experience of the operator?
- A simulation model of the water treatment could be beneficial for the operators (including a decision support system) but it should be different for each plant of EYDAP, as every plant is different.
- A what-if scenario which concerns EYDAP is if we upscale the DWT plant from the current situation, what could be the outcome in terms of disinfection products, disinfection by-products, energy consumption, dynamics and equipment etc.
- Total organic carbon is a substance in water body which is necessary for the operator to know online and to act on in terms of dosages and treatment efficacy.

From this brainstorm at EYDAP, it was thus concluded that within the ToDrinQ project soft-sensors should be developed mainly for the prediction of the source water quality to better operate current and future treatment systems.

3. Individual modules for monitoring and operational support

3.1 Introduction

Building on elements of past work (reaching Technology Readiness Level (TRL) 3) here we integrate real-time data and information on water quality from the entire system ('source to tap') obtained from online (soft and hard) sensors and support early warning, risk assessment and risk management. Advice on optimal control and risk management interventions throughout the water supply chain (source, conveyance, treatment and distribution), based on predictions of quality at the source and assessed variations in water quality during conveyance, treatment and distribution. The resulting platform can make of Nessie, a big data, FIWARE-compliant, dashboard as 'deployment route' for testing and validation by operators of the new methodologies and tools. Nessie is deployed as an online service for data management, analytics deployment, and visualisation and will allow ToDrinQ to have the important benefits of real end-user testing even for low TRL prototypes.

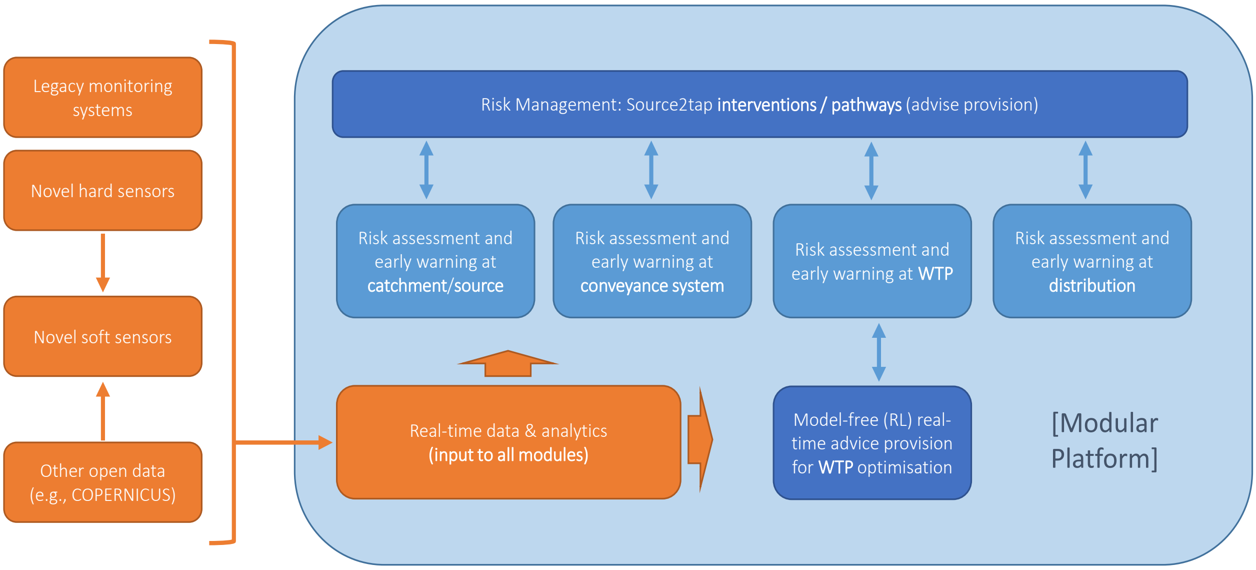


Figure 2: The integrated modular platform data and workflow

The new platform is expected to improve, in real-time, operations of water intake, conveyance, treatment and distribution, and support informed decisions when unexpected situations occur. This is especially important when the source is contaminated or large quality variations in the source occur, so fast risk mitigation is needed. We will develop a uniform risk-based approach (itself based on WSPs), integrating various, existing, relevant approaches. As shown in Figure 2 five main modules will be developed: 1. catchment/source, 2. Raw water conveyance network, 3. performance and state of DWT plant 4. distribution network 5. Source to tap intervention pathway design.

For each module (see Figure 3 and 4), information will be obtained from the (online) (soft or hard) sensor data (from WP3 and 4) and historical data to drive the model (available at the demo case e.g., SWIM for catchment/source module, EPANET-MSX for distribution network module etc, or simple but robust conceptual models if no case model exists). The outputs of the model feed the risk assessment sub-module to identify probabilistically long-term and short-term hazards, considering rules and thresholds set by Water Safety Plans (WSPs). Model outputs and probabilistic assessments are used as input to the downstream module (from source to tap) to identify water quality cascading effects. In parallel, the risk-assessment outputs are introduced in the cross-cutting source to tap intervention pathway module to provide advice on specific preventive and mitigation measures, related to long- and short-term hazards, respectively. Furthermore, short-term hazards trigger the early-warning sub-module to alert end-users to take actions. This model-agnostic procedure provides the appropriate flexibility to be implemented with different end-users' models and data, depending on the individual peculiarities of each premise. The WSP approach will underpin all modules and be based on both literature review and experiences of ToDrinQ's end-users (co-creation in collaboration with WP2).

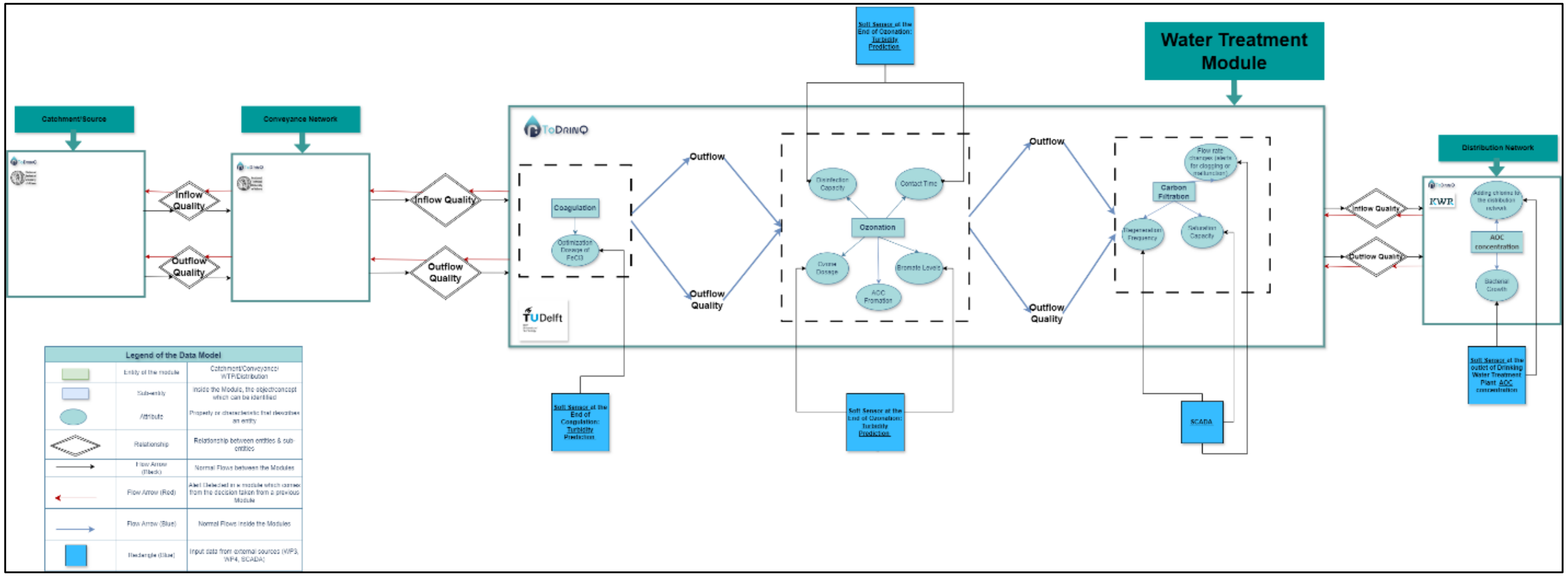


Figure 3: Water supply modules' representation expressed as Data Model

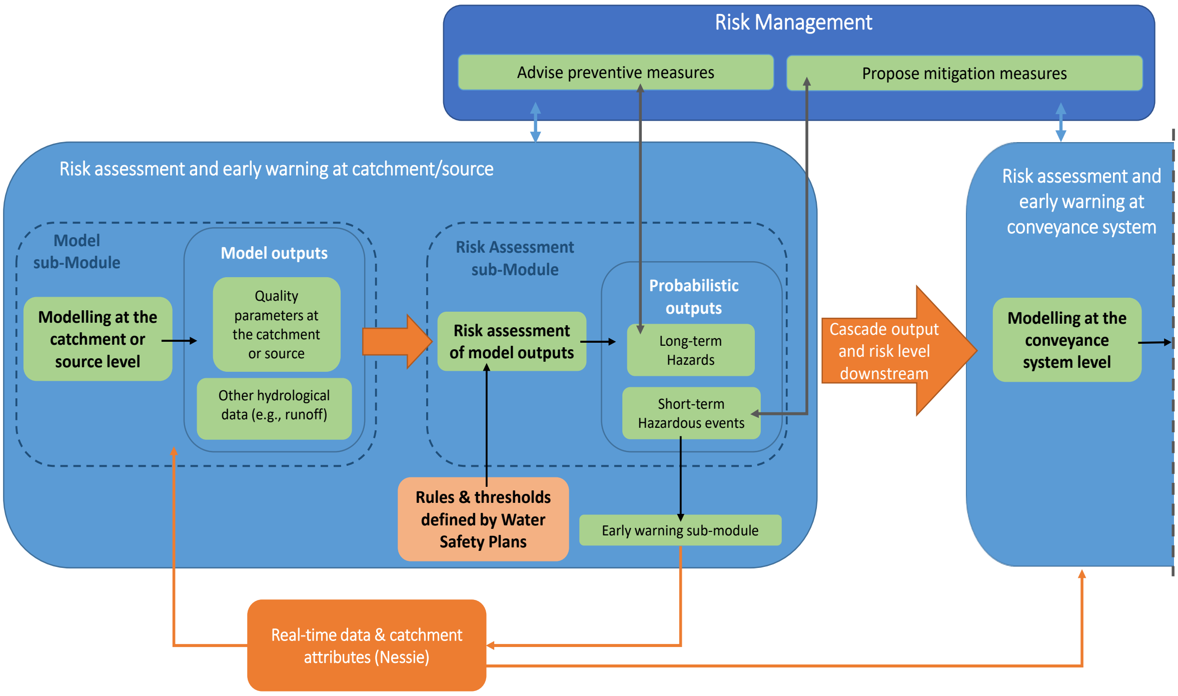


Figure 4: Looking inside the modules

Results from offline modelling can also be imported in the online modules through appropriate (FIWARE compliant) interfaces that will be defined in this work. This is important to allow different end-users to utilise their own models in combination with the ToDrinQ platform after the end of the project.

3.2 Risk-based Source to tap water supply management

From the Water Safety Plan (WSP)manual [1], it can be abstracted that "Effective planning for the supply of safe drinking-water must consider the growing uncertainties associated with a changing climate. Strengthening resilience can support water suppliers to better anticipate, respond to, cope with, recover quickly from, and adapt to, future shocks and stresses associated with climate variability and change. Water safety planning offers a systematic approach to build resilience to current and emerging climate threats by considering the implications of climate variability and change at each stage of the water supply. Water suppliers should consider past climate events that adversely affected the water supply and understand how projected changes in climate could threaten the system in the future. WSP teams may need to draw on external expertise, such as specialists in hydrology, climatology, public health and disaster risk reduction, to better understand the vulnerability of the system to the effects of climate change. Where required, system vulnerabilities must be addressed through robust improvement planning and strengthened management practices. Because climate projections are inherently uncertain, such measures should ideally provide benefit under different climate scenarios and be adaptable as new climate information becomes available."

---

[1] Water safety plan manual: step-by-step risk management for drinking-water suppliers, second edition. Geneva: World Health Organization; 2023. License: CC BY-NC-SA 3.0 IGO.

Based on the guidance from the United Nations (World Health Organization) and the EU institutions (EU Commission / DG Health, DG Environment), a risk assessment phase is fully integrated in the compliance process of the respective European legislation of the new (recast) Drinking Water Directive (EU) 2020/2184[2]. This mainly materializes through the WSP.

Water safety planning is a thorough approach to assess and manage risks across every stage of the water supply chain, from the catchment (source) to the tap (user). Recommended by the World Health Organization (WHO)[3], this method ensures the safety of drinking water and supports water suppliers in meeting quality standards and targets.

For this reason, the key features of the ToDrinQ modular platform have planned to integrate key elements (in a non-exhaustive way) from the water safety planning process according to European and international recommendations.

The water safety planning process can be summarized in the following steps:

- **Description of the water system** (answering the question on how the drinking water is being delivered from catchment to consumer). This includes the collection and documentation of the water system's information describing the entire supply chain. This step includes data collection from the field and from operators.
- **Identification of hazards and hazardous events** (identifying how something in the whole water supply chain could go wrong and that may affect the supply of safe drinking water). This phase builds upon the key elements already recognized during the detailed description. This is linked with Art 8 (2b) and Art 9 (2c). This step includes data collection from the field and from operators.
- **Validate existing control measures and assess risks and potential impacts** by combining the likelihood of occurrence and the severity of consequences of each of the hazards or compounds of them addressing their potential cascading effects.
- **Determine and validate control/mitigation measures and reassess or prioritize risks**

According to Art. 7 of the new (recast) Drinking Water Directive (EU) 2020/2184 risk-based approach to water safety is being introduced to cover the whole supply chain (starting from catchment area up to water consumption at tap/point of compliance).

The risk-based approach ensures that every stage of the water supply system is assessed for potential risks to water quality and public health. Even though water quantity issues are not explicitly mentioned, water quantity and quality are inextricably linked, especially during drought periods, flood events with impactful results in water quality. Water quantity dimension is prevalent mainly on catchment and conveyance elements of the water supply chain.

An initial identification of the taxonomy of a risk-based assessment and management conceptual framework, contributing to the purpose of ToDrinQ is presented in Table 2. Overall, we recognize 5 main dimensions (roots) for the risk assessment of a water supply system namely, physical components, risk and operational factors, regulatory compliance and stakeholder involvement (risk management and mitigation). Since many of the necessary inputs are based on existing EU environmental acquis, relevant columns have been added to ease further exploration of links and future developments.

---

[2] https://eur-lex.europa.eu/eli/dir/2020/2184/oj/eng

[3] https://openwho.org/courses/water-safety-planning

Table 2: Taxonomy of key elements of risk-based assessment containing physical components of the system, hazards and other risk factors as well as risk management and mitigation actors (regulation and stakeholders

Root	Tier 1	Tier 2	WFD Relevant	DWD relevant
Physical Components	Water Sources	Surface Water	X	X
		Groundwater	X	X
	Conveyance network	Natural	X	
		Artificial		
	Treatment Facilities	Filtration		
		Coagulation		
		Disinfection Units		X
	Distribution Network / Storage Facilities	Pipelines		
		Pumping Stations		
		Reservoirs		
		Water Tanks		X
	End Users	Residential		X
		Other Service sector		X
Hazards	Natural Hazards	Drought	X	
		Flood	X	
		Earthquake		
	Human-induced Hazards	Pollution	X	X
		Vandalism		
		Cyber-attacks		
Operational Factors	Operational Failures	Equipment Failure		
		Human Error		
	Supply and Demand	Water Demand Forecasting		
		Resource Allocation		
	Maintenance Activities	Scheduled Maintenance		
		Emergency Repairs		
	Monitoring Systems	Sensors and IoT Devices		
		SCADA Systems		
Regulatory Compliance	Quality Standards	Drinking Water Standards	X	X
	Legal Requirements	Environmental Regulations	X	X
	Reporting Obligations	Compliance Reporting	X	X
Stakeholders	Government Agencies		X	X
	Utility Companies		X	X
	Customers		X	X
	Environmental Groups		X	X

This taxonomical analysis could be further adjusted based on each of use case specific needs and might be used as driver to develop ontologies and (FIWARE-compliant) data models for each of the 4 modules. The way that the 5 physical components of each water supply systems are interlinked with the respective hazards and group risk management and mitigation strategies.

A summary Workflow for risk assessment and mitigation is as follows:

- Hazards Identification and Classification: Breakdown hazards for each element in the water supply chain (from catchment to tap).
- Dependency Mapping and Cascading Analysis: Model and assess interdependencies.
- Quantitative Risk Assessment: Use fault and event tree analysis, network models, and resilience metrics.
- Mitigation and Resilience Strategies: Focus on redundancy, robustness, and contingency plans.
- Monitoring and Early Warning: Implement real-time systems for proactive management.

With the **integrated modular platform** for monitoring and operational support ToDrinQ aims to give tools to operators and managers to anticipate on changes, e.g. provoked by climate change, integrating the expert knowledge and information on the system and to react in an adequate manner. Concerning the flow of the data into the risk assessment module and its architecture, the modular platform will host all the soft and hard sensors developed in ToDrinQ. In four out of five modules (Catchment/Source, Conveyance Network, Water Treatment, and Distribution Modules, as depicted in Figure 3), the prediction of the key parameters takes place. These predictions are given as input to the fifth module (Risk Assessment Module, as depicted in Figure 4). Apart from the prediction endpoints, for each of the soft sensors a table is provided with the risk, which is a multiplication of likelihood with severity determined based on expert knowledge and Drinking Water Directive thresholds.

First, concerning the likelihood numbers, they express the probability that a predicted parameter, generated by soft sensors, will exceed a predefined operational threshold. This probability is translated into likelihood grouped into three operational classes to facilitate consistent interpretation across all soft sensors. These classes represent how probable it is that a predicted parameter will exceed its acceptable threshold.

Table 3: Probability is translated to likelihood. There are three possible classes: low, moderate and high.

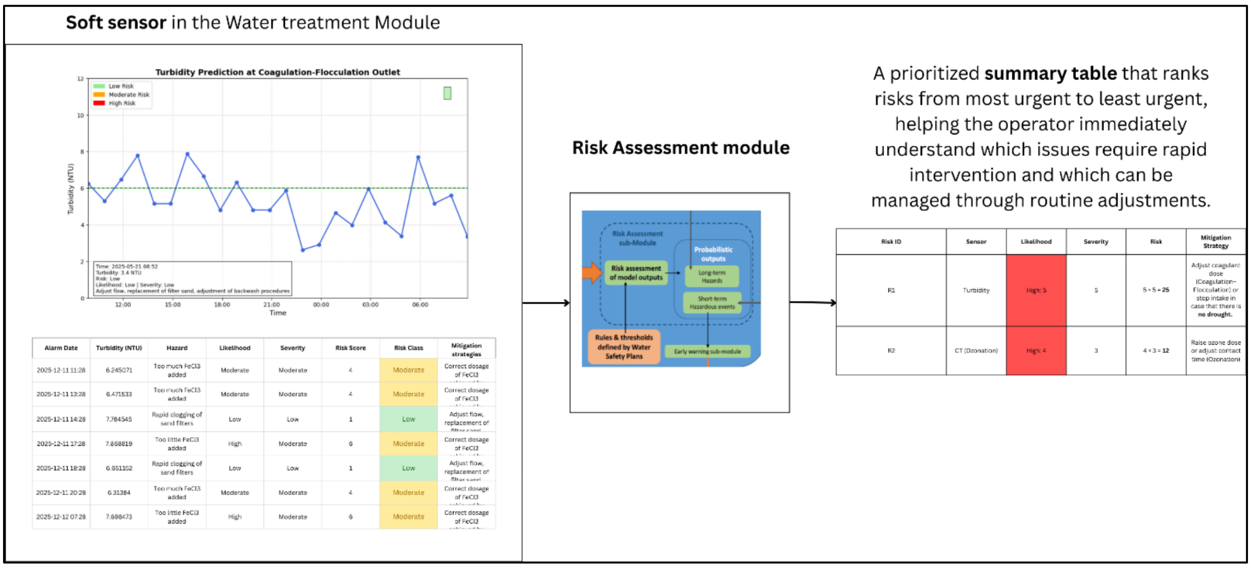
Likelihood class	Description	Approximate probability (p)	Operational meaning
Low (1–2)	Event is unlikely	($p < 0.15$)	Normal and stable operation; no intervention needed.
Moderate (3)	Event could occur under specific conditions	(0.15 $p < 0.40$)	Early signs of change; preventive measures recommended.
High (4–5)	Event is likely or imminent	($p > 0.40$)	High chance of threshold exceedance; corrective action required.

Second, concerning the severity, it expresses the **potential consequences or impact** if the hazard occurs and the predicted parameter exceeds the operational threshold. Unlike likelihood, which reflects the probability of occurrence, severity quantifies the **magnitude of the effect**, meaning how significant the deviation is in terms of operational, environmental, or water quality consequences. It is evaluated on a **five-level scale (1–5)** as presented in the following table.

Severity class	Description	Operational Meaning
Low (1–2)	Minor deviation from the operational threshold (approx. 5–10% difference).	Insignificant or long-term impact; no immediate intervention required.
Moderate (3)	Noticeable deviation affecting process performance or water quality (approx. 10–40% difference).	Potential issue; further investigation or trend analysis recommended.
High (4–5)	Significant deviation from the operational threshold (greater than 40%).	Short-term corrective action required to prevent deterioration or non-compliance.

Each risk, calculated as the product of likelihood and severity, is classified into three possible classes: low, moderate, and high risk. The resulting risk level is displayed to the user, operator, and/or decision-maker together with the corresponding mitigation strategy. Risks are evaluated for each module individually, but are also presented in a prioritized table, ranking them from the most urgent (requiring short-term action) to the least urgent (addressed through long-term measures).

To better illustrate the flow of data between the four modules (*Source, Conveyance, Treatment,* and *Distribution*) and the fifth, the Risk Assessment Module, a conceptual example for one of the soft sensors is provided below. The example refers to Soft Sensor about the prediction of turbidity at the outlet of coagulation-flocculation (SoSe #8 and RL#5). As shown, in addition to the prediction graph, the corresponding risk table for this soft sensor is also linked to the risk assessment module. Within this module, the tables and predictions of all soft sensors are compiled and evaluated. Based on this information, the output of the Risk Assessment Module is a summary table presenting all identified risks, prioritized from the most urgent (short-term risks) to the less urgent (long-term risks). This summary provides a probabilistic assessment of each risk, supporting timely decision-making and targeted operational responses from source to tap.



Alarm Date	Turbidity (NTU)	Hazard	Likelihood	Severity	Risk Score	Risk Class	Mitigation strategies
2025-12-11 11:28	6.245071	Too much FeCl3 added	Moderate	Moderate	4	Moderate	Correct dosage of FeCl3
2025-12-11 13:28	6.471533	Too much FeCl3 added	Moderate	Moderate	4	Moderate	Correct dosage of FeCl3
2025-12-11 14:28	7.784545	Rapid clogging of sand filters	Low	Low	1	Low	Adjust flow, replacement of
2025-12-11 17:28	7.868819	Too little FeCl3 added	High	Moderate	6	Moderate	Correct dosage of FeCl3
2025-12-11 18:28	6.651162	Rapid clogging of sand filters	Low	Low	1	Low	Adjust flow, replacement of
2025-12-11 20:28	6.31384	Too much FeCl3 added	Moderate	Moderate	4	Moderate	Correct dosage of FeCl3
2025-12-12 07:28	7.698473	Too little FeCl3 added	High	Moderate	6	Moderate	Correct dosage of FeCl3

A prioritized **summary table** that ranks risks from most urgent to least urgent, helping the operator immediately understand which issues require rapid intervention and which can be managed through routine adjustments.

Risk ID	Sensor	Likelihood	Severity	Risk	Mitigation Strategy
R1	Turbidity	High: 5	5	5 × 5 = **25**	Adjust coagulant dose (Coagulation–Flocculation) or stop intake in case that there is **no drought.**
R2	CT (Ozonation)	High: 4	3	4 × 3 = **12**	Raise ozone dose or adjust contact time (Ozonation)

Figure 5: Input and Output of the Risk Assessment Module. An example using synthetic data for one of the soft sensors in the Water Treatment Module is given

The output of the risk assessment is the probabilistic assessment of all the risks and their prioritization, from the most urgent to the less urgent, from source to tap.

The outputs and the relevant probabilistic assessments are then used as input to the downstream module (in a source to tap manner) to identify water quality cascading effects, see Figure 4. Since risks cannot always be mitigated within one (sub-)module, in parallel, the risk-assessment outputs are also introduced in a "source to tap intervention pathway design" module, to provide advice on specific preventive and mitigation measures, related to long- and short-term hazards, respectively. The source to tap intervention pathway design module has as inputs system information and data streams (legacy, WP3 and WP4 data streams) and risk levels (from the modules), as described in the methodology section and proposes alternative intervention pathways (from source to tap), to minimize (health) risk to consumers and increase consumer satisfaction and trust (promoting drinking from the tap). The source to tap intervention pathway module will be based on decision tree analysis, see Table 4 and Figure 6. The resulting 'risk pathway' will identify and depict risks as a map of possible causes, hazards and consequences. In this context, hazards refer to changes to the water supply operational performance (e.g., water treatment process failure). These hazards can be created by upstream internal and external causes (e.g., inadequate operational procedures, water quality disturbances) and create downstream consequences for the customer and community (e.g., failure to meet regulation, sickness) and thus need intervention. Depending on the situation, the module will suggest interventions to one or several of the sub-systems, but these interventions are often interrelated: e.g., when the source water deteriorates, intervention in the source can be identified, but also interventions in the drinking water treatment may be needed. An optimal set of interventions, resulting in a so-called 'intervention pathway' will be determined and communicated to the operator. Furthermore, short-term hazards trigger the early-warning sub-module to alert end-users to take actions. This model-agnostic procedure provides the appropriate flexibility to be usable by different end-users (with their own models and data).

Table 4: Basis for decision tree analysis for risk-based interventions pathways.

Potential Risk	Monitoring Sensor (Hard/Soft)	Likelihood	Severity	Mitigation Strategies
High concentrations and insufficient removal of turbidity/suspended solids/algae	✓ Turbidity ✓ Chlorophyl A	high	Low	✓ Stop intake ✓ Increased dosage of coagulant ✓ Application of pre-oxidation ✓ More frequent back-washing of filters
High concentrations of inorganic pollutants - heavy metals and nitrate - and insufficient removal	✓ Heavy metals ✓ Ammonium ✓ Nitrate	Moderate	Low	✓ Stop intake ✓ Avoiding nutrient run-off in catchment ✓ Increased dosage of coagulant. ✓ Dosage of acid/base (neutralization)
High concentration of OMPs, insufficient removal of OMPs	✓ OMPs ✓ UV254	High	High	✓ Stop of intake from source ✓ Higher dosage of O_3/UVH_2O_2 More frequent regeneration of GAC
High concentration of pathogens, insufficient removal	✓ Micro-organisms ✓ Viable cell counts	High	High	✓ Higher dosage of $O_3/Cl_2/UVH_2O_2$

Potential Risk	Monitoring Sensor (Hard/Soft)	Likelihood	Severity	Mitigation Strategies
Formation and insufficient removal of Biodegradable Organic Matter (BDOC, AOC)	✓ (AOC) ✓ UV254 ✓ DOC ✓ Ozone dosage	Moderate	Moderate	✓ Increased dosage of coagulant Lower dosage of O_3 ✓ More frequent regeneration of GAC
Excessive regrowth in distribution	✓ (AOC/BFP) ✓ DOC ✓ UV absorbance	Moderate	Moderate	✓ Increased dosage of coagulant ✓ Low dosage of O_3 More frequent regeneration of GAC
Occurrence of pathogens in distribution network	✓ Micro-organisms ✓ Viable cell counts ✓ Chlorine residual	Moderate	High	✓ Information of public and advice ✓ Increased dosage of chlorine Increased dosage of O_3/UV
Excessive formation BDPs	✓ UV254 ✓ DOC ✓ Cl_2 ✓ Residence time	Moderate	Moderate	✓ Increased dosage of coagulant ✓ Lower dosage of O_3/Cl_2 ✓ Optimization contact time of disinfectants during distribution
Excessive Corrosion	✓ pH ✓ Ca ✓ Alkalinity ✓ Saturation Index	Moderate	Moderate	✓ Dosage of acid/base (neutralization to adjust pH) ✓ Optimization of the water hardness

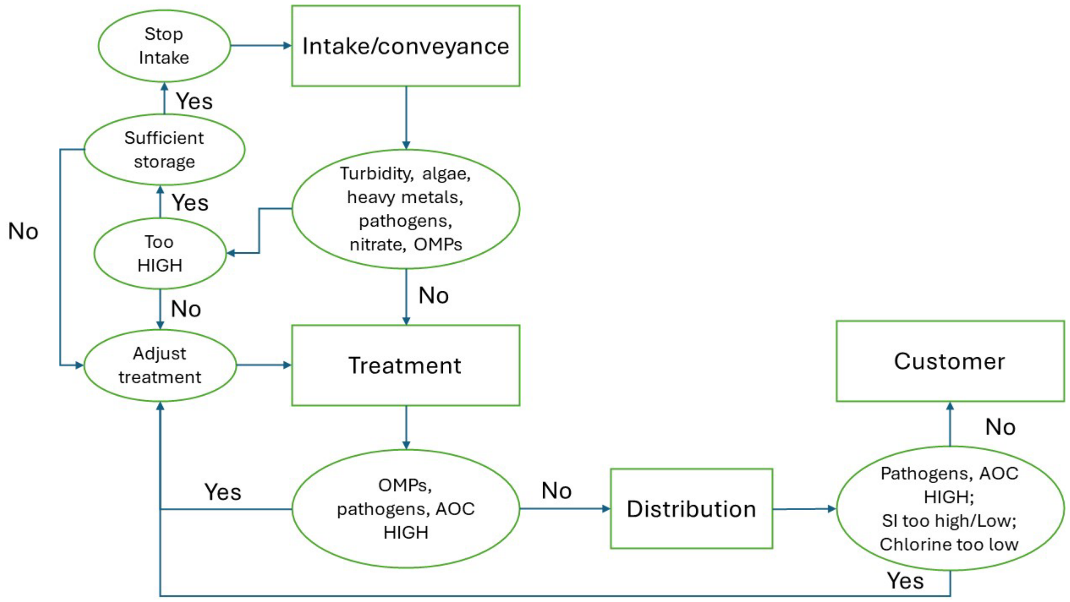


Figure 6: Example of decision tree for intervention pathways

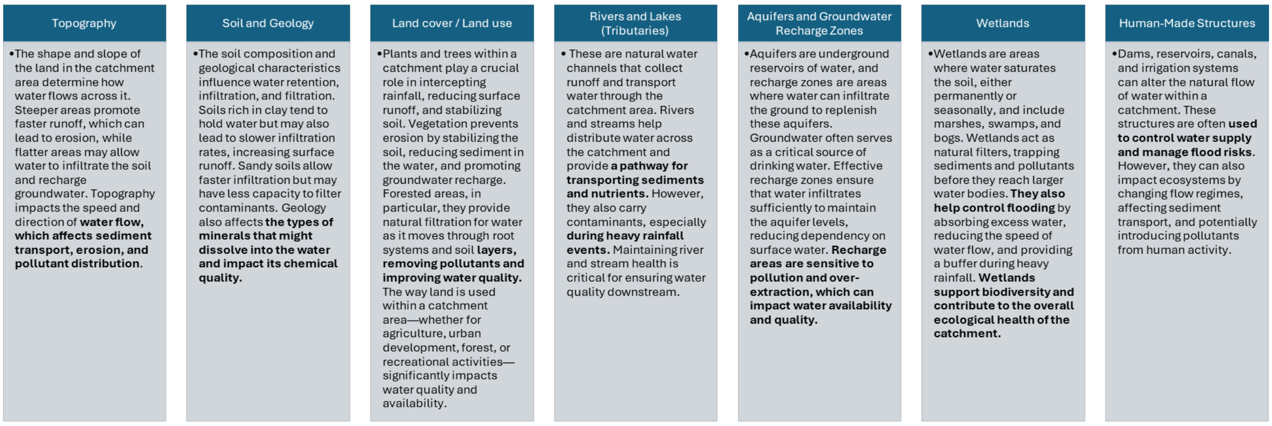
3.3 Catchment/source

The source or catchment element of a drinking water system is recognized as the initial barrier for safe water supply, encompassing the water source, land, and activities that may influence water quality before abstraction. Integrated treatment systems must consider the path the water travels upstream in the watershed. Renewed emphasis has been placed on source water protection as a critical strategy to minimize contamination risks, serve as a barrier for public health protection, and decrease treatment costs. A thorough system description, which includes topography, land uses, current sources (such as rivers, lakes, or groundwater), and potential alternative or emergency sources, is essential for identifying system vulnerabilities.

Within the ToDrinQ integrated platform, this protection strategy is operationalized through the Catchment/Source Module. This module moves beyond static description to dynamic monitoring by integrating Earth Observation data and Soft Sensors (developed in WP4) to assess water quality variations in real-time. Specifically, it focuses on early warning for critical hazards identified in the demo cases (such as nutrient runoff and algal blooms) and maps these dynamic inputs to the probabilistic risk assessment framework (Likelihood × Severity) established in Section 3.2.

This operational approach is applied to DC#2 (EYDAP), where Lake Yliki serves as a critical reservoir situated as the downstream water body at the outlet of the Boetikos Kifissos river basin. Due to this configuration, the lake acts as a sink for cumulative upstream pressures, necessitating specific monitoring of the Boetikos Kifissos inflows. Consequently, the module utilizes the Normalized Nutrient Load Index (NNLI) and specific soft sensors to monitor nutrient transport and runoff from the basin into the lake, directly addressing the agricultural hazards identified in the catchment.

Catchment elements that might be taken into consideration when elaborating a complete risk assessment framework are shown in Figure 7.



Topography	Soil and Geology	Land cover / Land use	Rivers and Lakes (Tributaries)	Aquifers and Groundwater Recharge Zones	Wetlands	Human-Made Structures
• The shape and slope of the land in the catchment area determine how water flows across it. Steeper areas promote faster runoff, which can lead to erosion, while flatter areas may allow water to infiltrate the soil and recharge groundwater. Topography impacts the speed and direction of **water flow, which affects sediment transport, erosion, and pollutant distribution**.	• The soil composition and geological characteristics influence water retention, infiltration, and filtration. Soils rich in clay tend to hold water but may also lead to slower infiltration rates, increasing surface runoff. Sandy soils allow faster infiltration but may have less capacity to filter contaminants. Geology also affects **the types of minerals that might dissolve into the water and impact its chemical quality.**	• Plants and trees within a catchment play a crucial role in intercepting rainfall, reducing surface runoff, and stabilizing soil. Vegetation prevents erosion by stabilizing the soil, reducing sediment in the water, and promoting groundwater recharge. Forested areas, in particular, they provide natural filtration for water as it moves through root systems and soil **layers, removing pollutants and improving water quality.** The way land is used within a catchment area—whether for agriculture, urban development, forest, or recreational activities—significantly impacts water quality and availability.	• These are natural water channels that collect runoff and transport water through the catchment area. Rivers and streams help distribute water across the catchment and provide **a pathway for transporting sediments and nutrients.** However, they also carry contaminants, especially **during heavy rainfall events.** Maintaining river and stream health is critical for ensuring water quality downstream.	• Aquifers are underground reservoirs of water, and recharge zones are areas where water can infiltrate the ground to replenish these aquifers. Groundwater often serves as a critical source of drinking water. Effective recharge zones ensure that water infiltrates sufficiently to maintain the aquifer levels, reducing dependency on surface water. **Recharge areas are sensitive to pollution and over-extraction, which can impact water availability and quality.**	• Wetlands are areas where water saturates the soil, either permanently or seasonally, and include marshes, swamps, and bogs. Wetlands act as natural filters, trapping sediments and pollutants before they reach larger water bodies. **They also help control flooding** by absorbing excess water, reducing the speed of water flow, and providing a buffer during heavy rainfall. **Wetlands support biodiversity and contribute to the overall ecological health of the catchment.**	• Dams, reservoirs, canals, and irrigation systems can alter the natural flow of water within a catchment. These structures are often **used to control water supply and manage flood risks**. However, they can also impact ecosystems by changing flow regimes, affecting sediment transport, and potentially introducing pollutants from human activity.

Figure 7: Key catchment elements of catchments characteristics to be considered

To ensure the Catchment/Source module is replicable across different European water systems, a universal mapping exercise was conducted to standardize the risk assessment logic. As illustrated in Figure 8, this framework systematically links physical Catchment Elements (such as topography, land cover, and aquifers) to potential Hazards (including drought, agricultural runoff, and flooding). These hazards are then mapped to their consequential Risks (e.g., nutrient loading, reduced water levels, or ecosystem disruption). This generalized taxonomy serves as the foundational ontology for the platform's risk algorithms, allowing the module to identify vulnerabilities regardless of the specific local context.

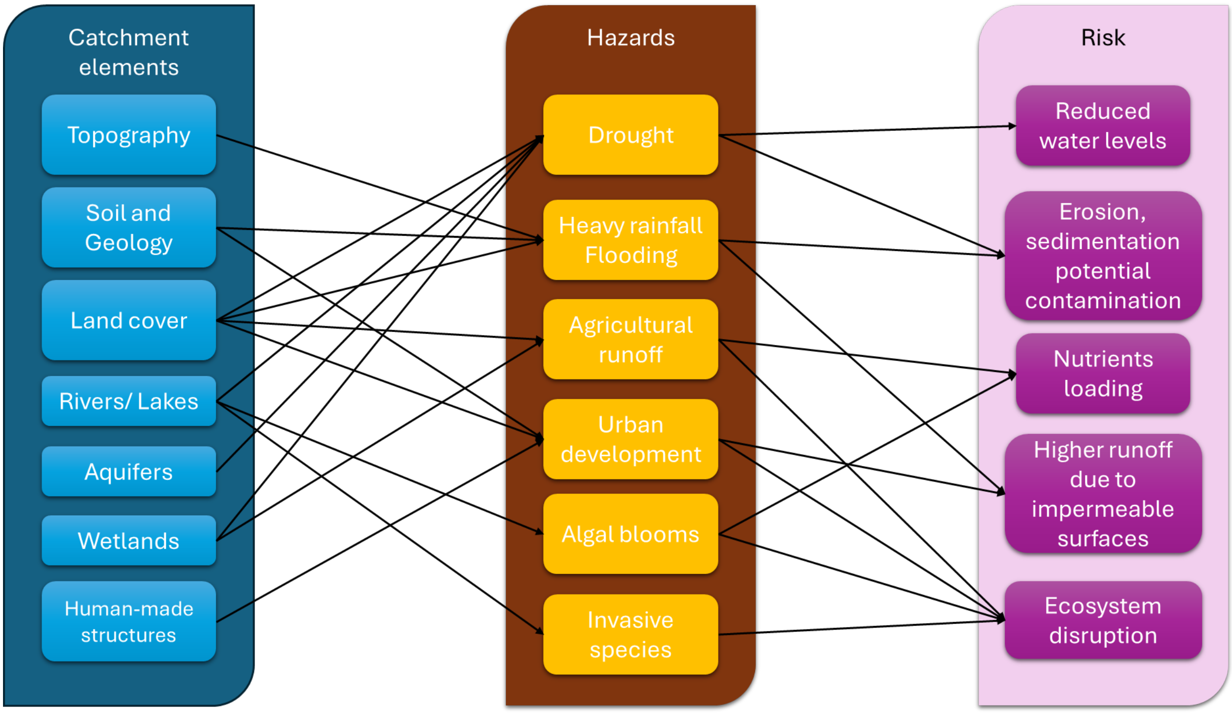


Figure 8: Mapping of catchment elements, potential hazards and potential risks

Operationalizing this risk framework for Lake Yliki, the module integrates three specific soft sensors developed and validated under WP4. These were selected based on their critical role in determining the lake's overall ecological status and treatment treatability:

- Chlorophyll-a (Chl-a) concentration (as a proxy for algal blooms/eutrophication),
- pH levels, and
- Dissolved Oxygen (DO) concentration.

A key innovation of these specific soft sensors is their reliance on Earth Observation (EO) data at their core. Unlike traditional in-situ sensors that provide data only at a single physical location, these EO-driven soft sensors generate estimations across the entire surface area of the lake. By exploiting this distributed nature of satellite data, **the module quantifies risk in a spatially explicit way**. This capability allows operators to detect not just when a threshold is breached, but exactly where in the lake the water quality degradation is occurring, enabling more targeted intake management and treatment adjustments.

Table 5: Hazard, risks and potential mitigation strategies for catchment quantity and quality issues

Hazard	Potential Risks	Data input from soft sensor	Mitigation Strategy	Likelihood	Severity
Agricultural nutrients runoff	Contaminant and nutrient loading cause water quality degradation (eutrophication).	Nutrient Runoff Soft Sensor: Monthly Normalized Nutrient Load Index (NNLI) combined with Daily Discharge Forecast.	Establish buffer zones, enforce runoff regulations, and adjust intake ratios based on load forecasts.	High	Major
Algal blooms	Water quality degradation, potential toxin release, and physical clogging of treatment plant screens.	Chl-a Soft Sensor: Spatially explicit estimation of Chlorophyll-a concentration (Critical threshold ≥8μg/l).	Stop intake from affected zones, increase coagulant dosage downstream , and use biological controls.	Moderate	Moderate
Hypoxia / Low Oxygen	Release of sediment-bound metals (e.g., Manganese, Iron) and ecosystem disruption due to anoxic conditions.	Dissolved Oxygen (DO) Soft Sensor: Spatially explicit estimation of DO concentration (Critical threshold ≤8mg/l).	Adjust intake depth to avoid anoxic layers; prepare for increased oxidation requirements at the treatment plant.	High	Major
Acidification	Corrosion of conveyance infrastructure and disruption of chemical treatment processes downstream.	pH Soft Sensor: Spatially explicit estimation of pH levels (Critical threshold ≤7).	Dosage of base (neutralization) at treatment inlet; monitoring for acidic event trends.	Low to Moderate	Moderate

For DC#2, the prediction of the turbidity in the source was identified as critical for the performance of the treatment plant.

The risk assessment modules development process at source scale are described below. These modules utilize the results of the soft sensors for the monitoring of water quality variables at source scale. outputs of the risk assessment modules will be further described in Deliverable 7.2.

3.3.1 Nutrient runoff

The risk assessment module, developed for DC#2, follows the development of the Early Warning System for nutrient runoff as described in Deliverable D4.3. The outputs of this soft sensor include the daily discharge forecast and the monthly nutrient load estimation at the basin, using the proposed Normalized Nutrient Load Index (NNLI). This module analyses these two variables, and thus two risk classification methods were applied.

Classification of Daily Discharge

Daily discharge was classified into four groups, based on its tole as a nutrient transport mechanism. The analysis was performed using data from the 28-year period between 1994 and 2022. The classification applied is presented in Table 6.

Table 6: Discharge Classification

Discharge Class	Hydrological Condition	Classification Rule	Hydrological / Ecological Function
Zero	No Inflow	Discharge = 0	No nutrient transport
High	High Flow / Flushing	Discharge > 66.7th percentile (0.776 hm^3)	High nutrient transport
Medium	Medium Flow	Discharge > 33.3rd percentile (0.195 hm^3)	Moderate flow – Moderate transport
Low	Low Flow	Discharge > 0 or Discharge < 33.3rd (0.195 hm^3)	Low nutrient transport but high concentration risk

Nutrient Load Classification

The nutrient load classification was calculated based on the Normalized Nutrient Load Index (NNLI). This index is calculated monthly, based on the agricultural activities required by the basin’s crops (see Deliverable 4.3). The monthly NNLI ranges from 0 to 1 and the value representing each month’s nutrient load has been calculated. The NNLI monthly values were split into three groups (see Table 7):

Table 7: Monthly Nutrient Load Classification

Nutrient Class	NNLI Range	Months	Condition
Good	0.00-0.054	Aug, Sep, Jan	Low nutrient load at the basin
Medium	0.189-0.379	Feb, May, Jun, Jul, Oct, Nov, Dec	Medium nutrient load at the basin
Bad	0.539	Mar, Apr	High nutrient load at the basin

Risk Classification

A matrix was used to create the final risk classification for nutrient discharge in the lake. The risk assessment will be conducted daily, each time a new discharge forecast becomes available. The approach described below combines the Nutrient Load Class (Table 7) and the Discharge Class (Table 6) to define the overall nutrient discharge risk, assigning bands of R1(Low Risk), R2(Medium Risk) and R3(High Risk), provided at Table 8.

- The concentration problem, which is defined as low discharge combined with “high” nutrient load (“Bad” nutrient conditions).
- The flushing problem which is designed as “high” discharge (high transport) combined with medium or high nutrient load (“Medium” or “Bad” nutrient conditions).

Table 8: Combined Risk Assessment Matrix

	Nutrient: Low Index Month	Nutrient: Medium Index Month	Nutrient: High Index Month
Discharge: Zero	**R1**	**R1**	**R1**
Discharge: Low	**R1**	**R2**	**R3**
Discharge: Medium	**R1**	**R2**	**R3**
Discharge: High	**R2**	**R3**	**R3**

3.3.2 Concentrations of Chl-a, pH and DO

The process followed for the development of the risk assessment modules for Chl-a, DO and pH are very similar and use as inputs the maps generated by the soft-sensors described in Deliverable 4.3, as developed for DC#2. To establish the baseline for these risk assessment modules, we utilized the generated historical data spanning from 2020 to 2025 of Chl-a concentration, DO concentration and pH maps. For the efficient spatiotemporal analysis these maps were aggregated and transformed into unified Image Time Series Data Cube that allows for pixel querying along the spatial and temporal dimension.

Methodology and Threshold Definition

The risk assessment module development is driven by expert judgement regarding the critical levels of these three water quality values. The methodology is based on two threshold levels:

- **The value threshold (Hazard)**: The value of the water quality variable (Chl-a and DO concentration and pH values): This threshold sets the critical threshold. This is the threshold above or below which the water quality is considered compromised.
- **Spatial Extend Threshold (Severity):** For every time instance in the DataCube, the risk assessment module developed calculates the percentage of the total valid pixels in the lake surface area where the pixel value exceeds or is lower than the critical threshold.

More specifically, the critical value threshold for Chl-a is set to values **>= 8µg/l**, for DO is set to **<= 8mg/l** and for pH the threshold is set to **<= 7** (see Table 9).

Table 9: Critical thresholds for Chl-a, pH and DO

Variable	Critical Threshold	Comment
Chl-a	>= 8µg/l	Expert defined threshold – risk of eutrophication
DO	<= 8µg/l	Conservative threshold – Yliki is well oxygenated
pH	<= 7	Detect potential acidic events

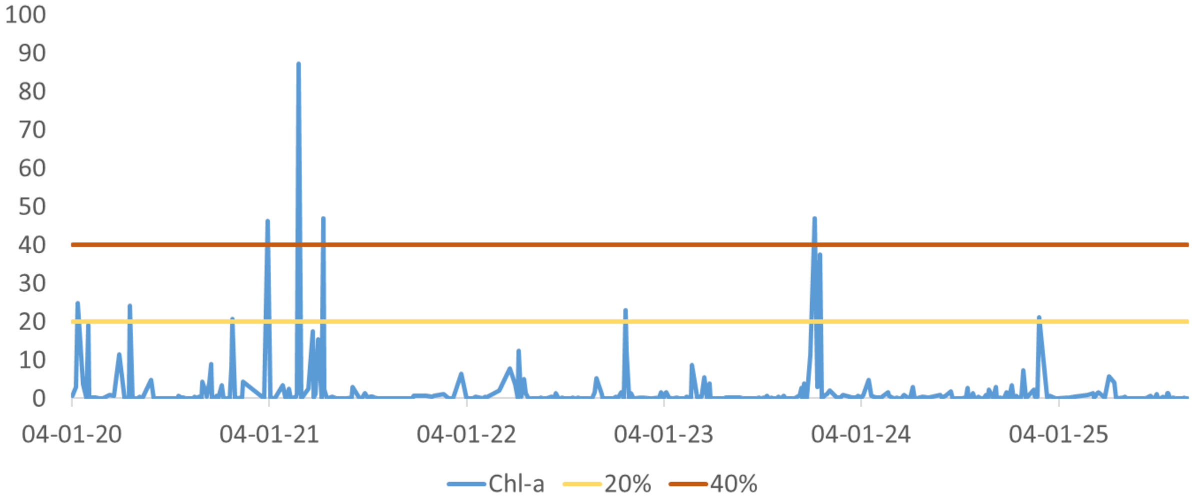
Operational Risk Classification

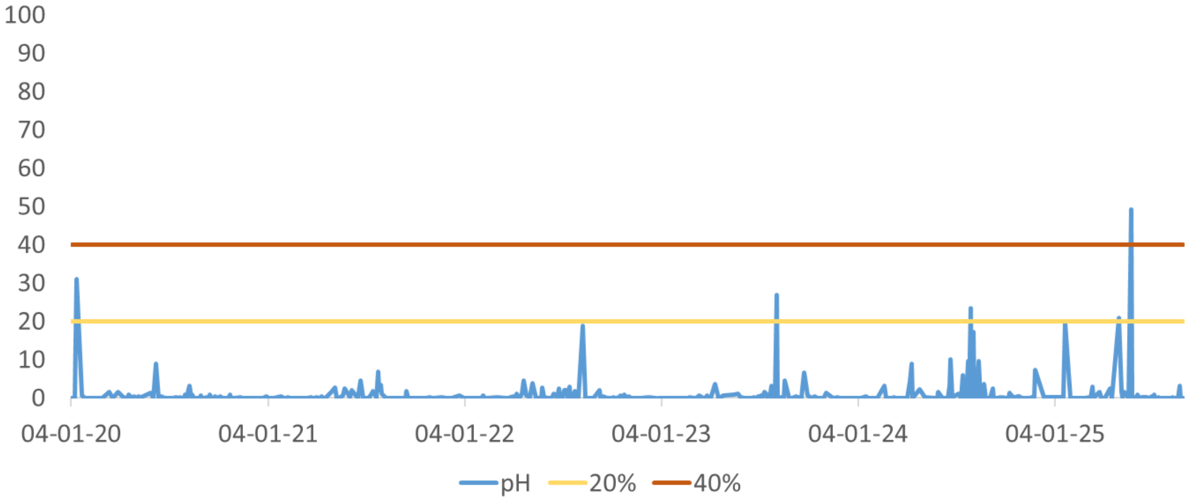
At the operational level, each new soft-sensor output (Chl-a, DO and pH) is processed against the historical analysis described above. The module calculates the spatial extent of the exceedance for each new instance and assigns a risk category based on the defined ranges provided below (see Table 10).

Table 10: Risk Classification Thresholds

Risk Classification	Spatial Extend of Exceeding Pixels	Short Description
R1: Low Risk	0% - 20%	Small area of the lake exceeds (or is below) the safety threshold
R2: Medium Risk	20% - 40%	A larger but not extreme area of the lake is affected
R3: High Risk	>40 %	A big area of the lake exceeds (or is below) the critical threshold

The following figures provide the time series of the pixels exceeding the critical thresholds and the thresholds of their spatial extend for each of the three variables (see Figure 9). For the DO the threshold is only one and is set to 40%. Above this threshold the risk is characterized as medium (R2) because no serious hypoxic events were identified, and the value threshold for this variable is moderate.





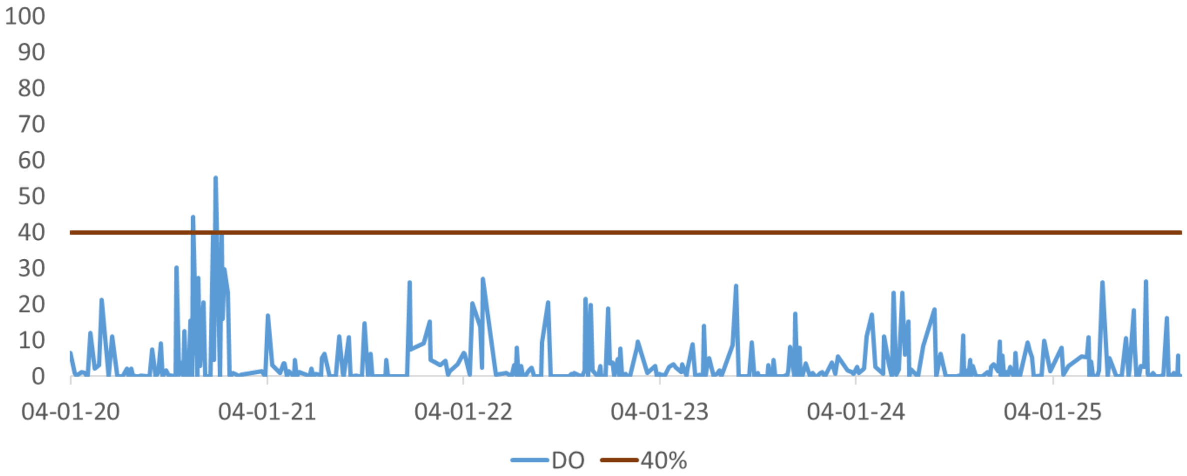


Figure 9: Time series of exceeding pixels and corresponding thresholds

Finally, this automated classification is suitable for immediate risk assessment and decision making as soon as the new soft-sensor outputs become available.

3.3.3 Turbidity at the source

Within the source water module of DC#1 –(WTNT), a raw water turbidity soft sensor is included to provide an early indication of deteriorating source water conditions before water enters the treatment plant. The objective of this soft sensor is to support proactive operational decision-making by predicting changes in source water quality that may affect downstream treatment processes and ultimately increase the risk of not meeting drinking water quality requirements.

Raw water turbidity is considered a key operational indicator because elevated turbidity levels are often associated with increased particulate matter, organic loading, and variations in source water quality. These changes may affect the performance of subsequent treatment processes, particularly coagulation-flocculation and filtration. By forecasting turbidity several hours ahead, operators are provided with additional response time to prepare treatment operations and optimize process settings.

Furthermore, knowledge of incoming turbidity conditions enables more efficient operation of the coagulation process. Since ferric chloride ($FeCl_3$) dosage is closely related to the particulate load entering the plant, early turbidity predictions can support operators in adjusting coagulant dosing strategies, improving contaminant removal efficiency while avoiding unnecessary chemical consumption and associated operational costs.

Therefore, the predicted turbidity is subsequently linked to the Treatment module and specifically with the SoSe #8: Coagulation performance model which serves a performance indicator for the effectiveness of the coagulation process.

3.4 Conveyance network

The conveyance system, which includes critical infrastructure such as intakes, transmission systems, and aqueducts, is a basis of the overall risk assessment process in water safety planning. Within the comprehensive, proactive risk assessment framework of Water Safety Plans (WSP), every stage of the water supply must be assessed for hazards and hazardous events. Consequently, security vulnerability assessments are mandatory for critical assets like transmission systems, and the fundamental design philosophy is that the failure of any single component (such as an element of pipe or a valve) should not put the entire facility out of service.

To protect this critical infrastructure, a key focus is placed on reliability and redundancy, achieved through practices such as designing intakes for operation under adverse conditions and constructing duplicate intake structures or parallel conduits for major facilities. Key parameters and protective measures to watch include maintaining hydraulic integrity by designing intake conduits to minimize sediment deposition and ensuring systems operate with positive pressure to prevent back-siphonage contamination. Physical protective systems involving barriers, secured hatches, and locks are vital, complemented by monitoring mechanisms. Strategic implementation and deployment of online sensors track critical water quality indicators (such as pH, turbidity, and specific conductivity) to detect potential entry of contaminants during transport.

Within the ToDrinQ platform, the Conveyance Network Module has been designed to assess risks and provide early warnings for water quality incidents occurring during transport (specifically focusing on turbidity events and conductivity spikes) that could overwhelm downstream treatment plants. Unlike the Catchment Module which relies heavily on Earth Observation, this module is primarily driven by the integration of legacy data. This integration allows for the detection of cascading risks where physical infrastructure failures (e.g., pump degradation) directly impact water quality (e.g., sediment mobilization).

This module has been configured for DC##2 (EYDAP), managing the raw water transport from the Yliki aqueduct. Based on the user needs analysis (Section 2.2), a critical operational challenge identified by EYDAP was the intrusion of sand and high turbidity, particularly when mixing groundwater from boreholes with surface water. The Conveyance Module directly addresses this by assessing turbidity and conductivity, providing alerts to protect pumping station impellers and prevent filter clogging at the Polydendri Drinking Water Treatment Plant. The key elements of conveyance systems are presented in Figure 10 below.

Pipelines and aqueducts	Pumping Stations	Valves and Flow Control Devices	Monitoring and Sensor Systems	Water Intakes
• Primary conduits for moving water over both short and long distances, usually underground or occasionally above ground in certain terrains. Pipelines are the main transport routes and must withstand pressure changes, environmental stresses, and potential corrosion.	• Pumping stations provide the necessary pressure to move water through pipelines, especially in regions with varied topography or where water needs to travel long distances. Pumps are critical for overcoming gravity in uphill conveyance and maintaining consistent water pressure. Redundancy and backup power are essential to ensure reliability during power outages or equipment failures.	• Valves regulate water flow, pressure, and direction within the conveyance network. Valves help manage flow rates, prevent damage from water hammer (sudden pressure surges), and allow for isolation of sections during maintenance, protecting the system from failure.	• Real-time monitoring allows for early detection of problems, such as leaks, pressure drops, or contamination, enabling quick responses to prevent or mitigate damage and service disruptions.	• Water intakes are the entry points for water entering the conveyance system, often equipped with screens and filters to remove debris. Proper intake design prevents large debris and contaminants from entering the system, reducing the burden on conveyance and treatment facilities and protecting pipelines and pumps from damage.

Figure 10: Key catchment elements of conveyance systems characteristics to be considered

Following up on the risk-based assessment framework, we mapped – in a non-exhaustive way- potential links of catchment elements, hazards and risk as shown in Figure 11 below.

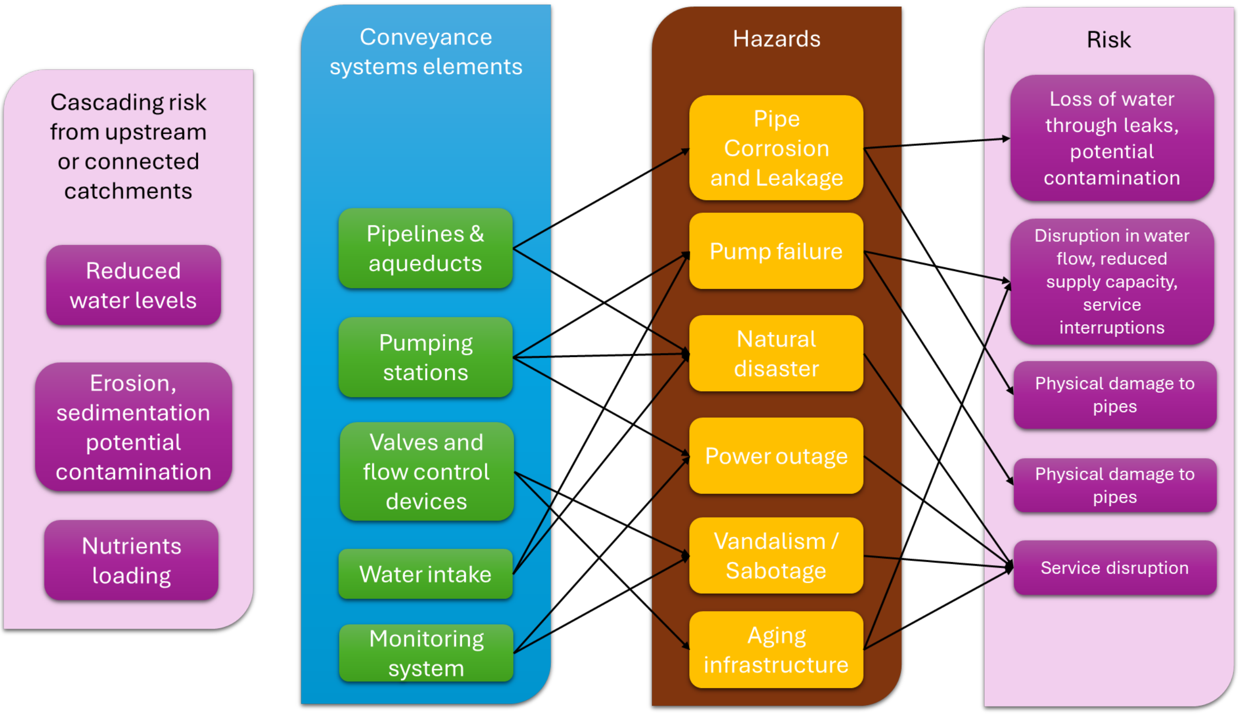


Figure 11: Mapping of conveyance elements, potential hazards and potential risks

The definition of "recommended" operational ranges in the Greek context cannot be decoupled from the mandatory legislative framework. The operational envelopes for turbidity and conductivity are legally bounded by statutes that transpose European Union directives into national law, creating a rigid hierarchy of water quality categories that dictate the permissible methods of abstraction and treatment.

The cornerstone of raw water quality regulation in Greece is the Joint Ministerial Decision (KYA) 46399/1352/1986, titled "Quality required of surface water intended for the abstraction of drinking water[4]". This legislation was the national transposition of Council Directive 75/440/EEC. While subsequent directives such as the Water Framework Directive [5](2000/60/EC) have broadened the scope of water management to include ecological status, KYA 46399/1352/1986 remains the specific operational benchmark for defining the suitability of surface water for potable use based on physicochemical loads. The legislation classifies surface water sources into three distinct categories: A1, A2, and A3, based on the intensity of treatment required to render the water potable. This classification system is fundamental to operational planning, as it sets the "ceiling" for parameters like conductivity and turbidity beyond which water is deemed unfit for abstraction without advanced, often economically unviable, treatment.

Reflecting the critical operational priorities identified during the co-creation phase with EYDAP, the Conveyance Network Module operationalizes risk assessment for two primary hazards, sediment intrusion and water source mixing, as detailed in the following Table 11.

---

[4] Required quality of surface waters intended for drinking purposes - link

[5] DIRECTIVE 2000/60/EC - link

Table 11: Hazard, risks and potential mitigation strategies for conveyance quantity and quality issues

Hazard	Potential Risks	Data Input from legacy data collection system	Mitigation Strategy	Likelihood	Severity
Sand / Sediment Intrusion	Physical damage to pump impellers; rapid clogging of treatment plant intake screens due to high suspended solids	Turbidity / Suspended Solids (Online) Three risk bands identified based on turbidity levels 0-10 NTU 10-50 NTU >50 NTU	Reduce pumping rate during high-sediment events; activate pre-sedimentation protocols; switch intake source to avoid peak turbidity loads.	High	Major
Water Source Mixing / Hardness Spikes	Deviation from water quality targets due to high hardness or salinity; potential contaminant ingress detected via ionic strength changes .	Conductivity (Online)	Adjust mixing ratios of groundwater (borehole) and surface water; isolate specific boreholes with high conductivity values; alert DWTP to adjust softening processes.	Moderate	Moderate

3.4.1 Turbidity risk assessment

Currently there are no specific recommended operational ranges for turbidity for raw water conveyance systems in Greece. The operational ranges in raw water conveyance systems are typically determined by two factors: the characteristics of the raw water source (here in DC#2 is usually a 70-30% mix of surface and groundwater respectively) and the requirements of the downstream water treatment plant (at Polydendri), as the goal of an intake system is to reliably deliver water of the best available quality. Turbidity represents on the most volatile operational parameters raw water conveyance systems. It is driven by stochastic meteorological events (e.g. heavy rainfall, flash floods, and landslides) and dictates the immediate treatment strategy of the water supply.

The dataset (spanning from July 2024to October 2025) was segmented into three risk categories based on the proposed turbidity thresholds: Low (<10 NTU), Medium (10–50 NTU), and High (>50 NTU) as shown in Table 12. This classification reveals that the raw water remains within the low-risk "normal" operating range for approximately 82% of the recorded period. The medium band, representing a transitional or alert state, accounts for nearly 12% of observations, while high-turbidity events requiring significant intervention occur approximately 6% of the time.

Table 12: Turbidity risk bands

Risk Level	Turbidity Range (NTU)	# of records	Percentage of values in this range
R1: Low	**< 10**	7910	**82.03%**
R2: Medium	**10 – 50**	1123	**11.65%**
R3: High	**> 50**	610	**6.33%**

By sorting the turbidity values in ascending order, Figure 12 provides a clear visual representation of the cumulative data distribution, effectively highlighting the proportion of operational time spent within each defined risk band.

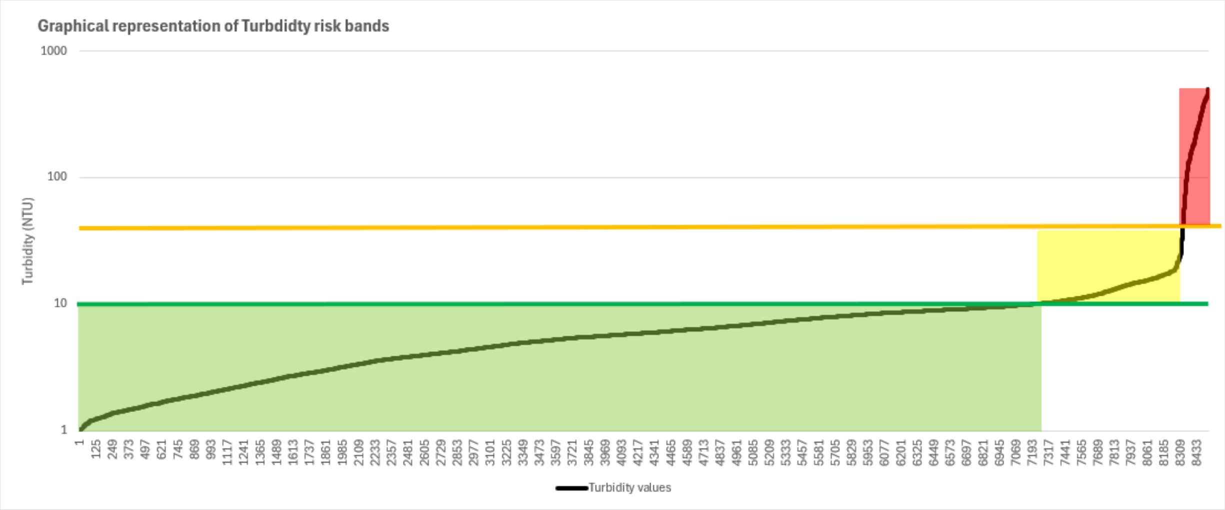


Figure 12: Risk bands as indicated by the data analysis and operational turbidity levels

In Figure 13 we present the temporal variation of raw water turbidity dataset, overlaid with the expert-defined risk bands to visualize operational stability and event severity. The Low-risk (R1) zone (green, <10 NTU) captures the baseline operating conditions, while the Medium-risk (R2) band (yellow, 10–50 NTU) highlights transitional periods and minor fluctuations. The High-risk 9R3) zone (red, >50 NTU) clearly determines significant turbidity spikes, allowing for quick identification of critical events where water quality exceeds standard treatment thresholds.

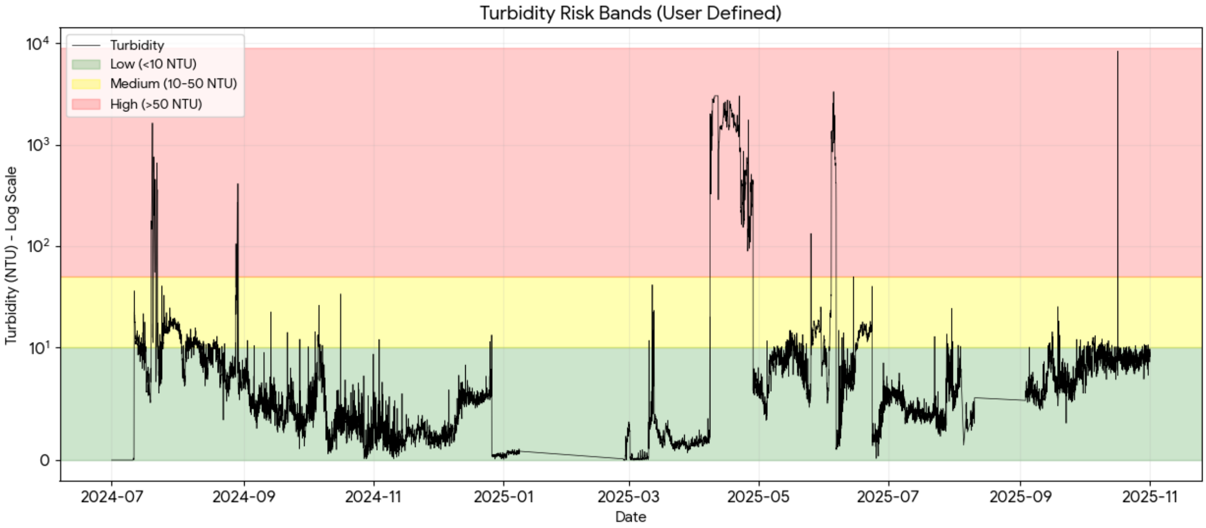


Figure 13: Turbidity risk bands projected in the available time series

3.4.2 Conductivity risk assessment

Currently there are no specific recommended operational ranges for conductivity for raw water conveyance systems in Greece. Electrical conductivity of raw water conveyance is a parameter of stability. It reflects the geochemical signature of the catchment area and serves as a sentinel for systemic changes. The operational management of conductivity is less about immediate "treatment" (as dissolved ions are difficult to remove) and more about source selection and infrastructure protection.

The analysis classifies the raw water conductivity data into three distinct risk bands: Low (<400 µS/cm), Medium (400–450 µS/cm), and High (>450 µS/cm). The data indicates that the system operates within the stable, low-risk baseline for approximately 80% of the time. However, 12% of the recorded period reflects an intermediate alert state with elevated dissolved solids, while the remaining 8% represents critical peak intrusion events where conductivity exceeds 450 µS/cm, signalling a significant deviation from normal water quality parameters.

Table 13: Conductivity risk bands

Risk Level	Turbidity Range (µS/cm)	# of records	Percentage of values in this range
R1: Low	**< 400**	7689	**79.74%**
R2: Medium	**400 – 450**	1175	**12.19%**
R3: High	**> 450**	779	**8.08%**

By sorting the conductivity values in ascending order, Figure 14 provides a clear visual representation of the cumulative data distribution, effectively highlighting the proportion of operational time spent within each defined risk band.

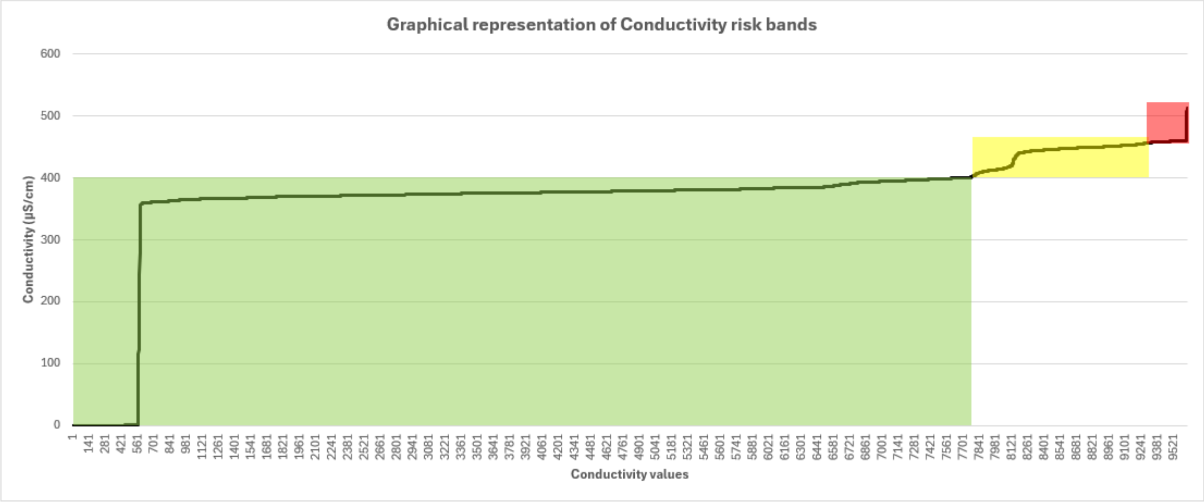


Figure 14: Risk bands as indicated by the data analysis and operational turbidity levels

The updated graph illustrates the temporal fluctuation of raw water conductivity, overlaid with color-coded risk bands to distinguish operational states. The Low-risk zone (green, <400 µS/cm) encapsulates the stable baseline where the system operates the majority of the time. The Medium risk band (yellow, 400–450 µS/cm) highlights transitional periods of elevated conductivity, while the High-risk zone (red, >450 µS/cm) clearly marks distinct peak events, allowing for visual identification of critical periods where water quality significantly deviates from the norm.

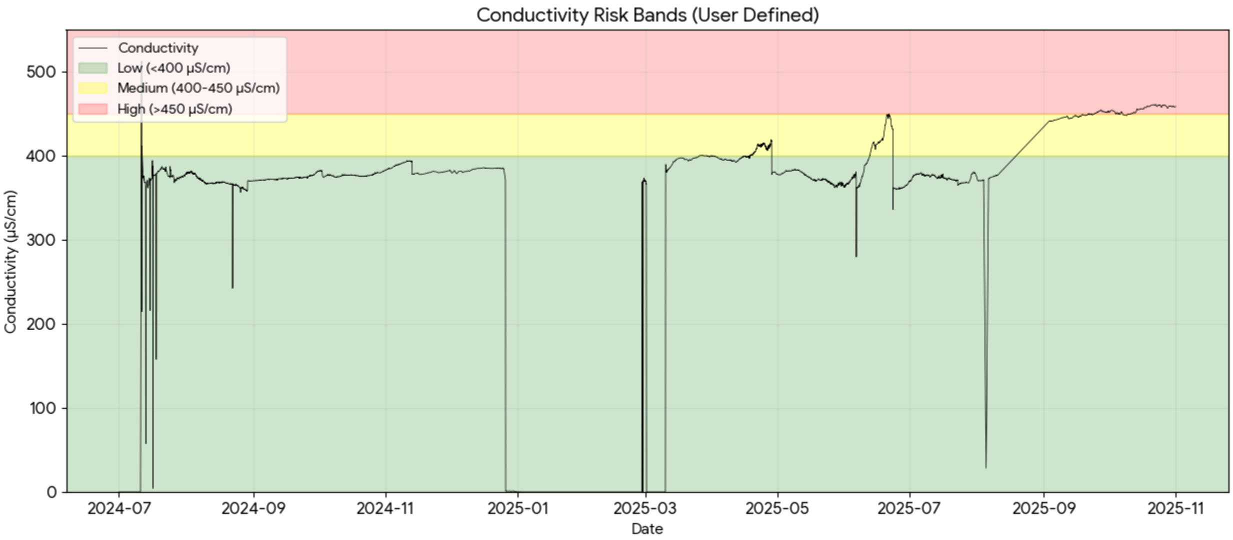


Figure 15: Conductivity risk bands projected in the available time series

3.5 Performance and state of water treatment

The module on performance and state of the water treatment plant assesses risk of not meeting the desired drinking water quality also in relation to the state of the system e.g., by identifying the difference in performance between clogged and clean filters and optimal cleaning procedures.

For the water treatment module, we therefore defined the following set of upstream effects:

- Increasing concentration of micropollutants in the inlet water
- Increasing particulate matter in inlet water
- Fluctuating bacterial or viral load in inlet water
- Increasing natural organic matter in inlet water
- Fluctuation of environmental parameters (pH, temperature, salt concentration) of inlet water

We defined the following treatment elements:

- Coagulation-flocculation(-sedimentation)
- Rapid Sand filtration
- Ozonation
- Hardness reduction/softening
- Activated carbon filtration

In Table 14, a non-exhaustive list of risks associated with these treatment technologies are given.

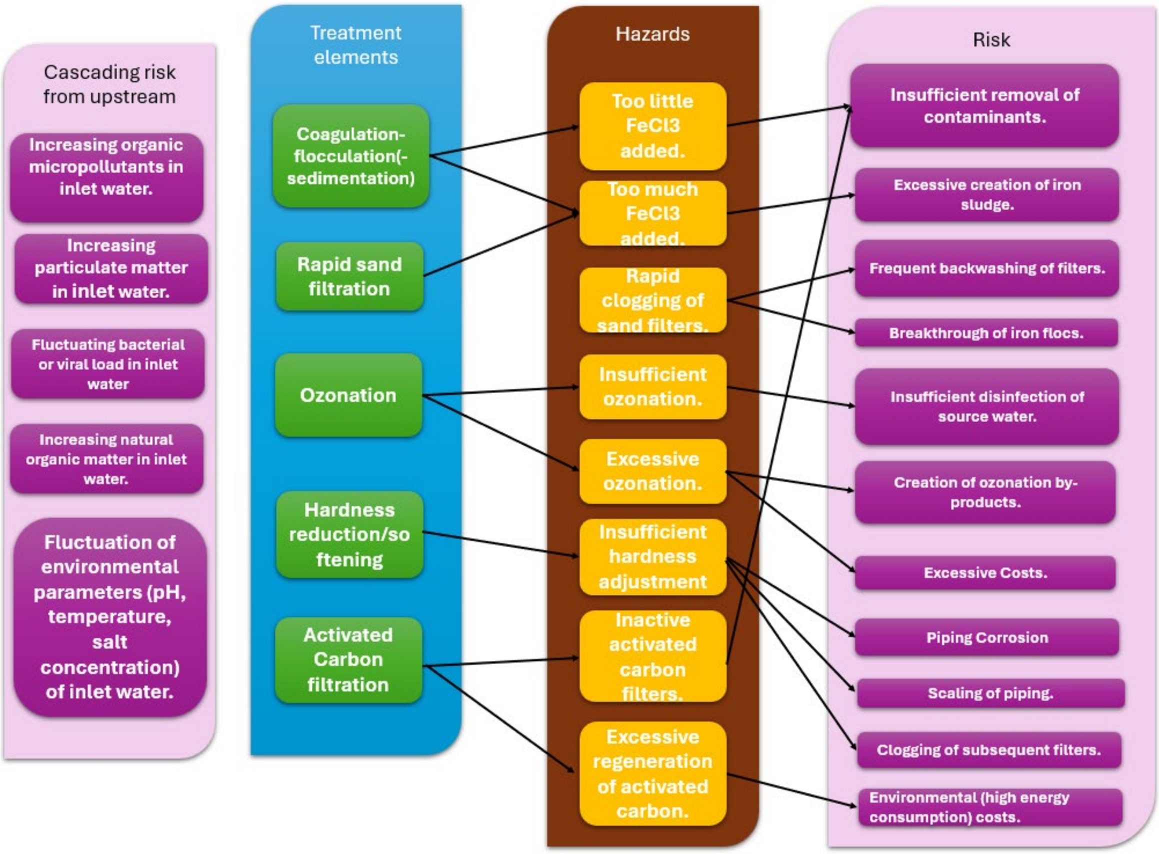


Figure 16: Mapping of water treatment elements, potential hazards and potential risks

Table 14: Hazard, risks and potential mitigation strategies for treatment issues

Hazard	Potential Risks	Mitigation Strategy	Likelihood	Severity
Too little $FeCl_3$ added	Insufficient removal of contaminants such as particulate material and organic matter	Correct dosage of $FeCl_3$ achieved by constant monitoring of output turbidity	High	Moderate
Too much $FeCl_3$ added	Excessive creation of iron sludge, possible insufficient removal of contaminants	Correct dosage of $FeCl_3$ achieved by constant monitoring of output turbidity	Moderate	Moderate
Rapid clogging of sand filters	Frequent backwashing of filters, breakthrough of iron flocs.	Adjust flow, replacement of filter sand, adjustment of backwash procedures	Low	Low
Insufficient ozonation	Insufficient disinfection of source water	Constant monitoring of AOC, bromate and correct contact time ozone and water	High	Major
Excessive ozonation	Creation of ozonation by-products which are detrimental to health, excessive costs	Constant monitoring of AOC, bromate and correct contact time ozone and water	Moderate	Moderate

Hazard	Potential Risks	Mitigation Strategy	Likelihood	Severity
Insufficient hardness adjustment	Corrosion or scaling of piping Rapid clogging of subsequent filters	Adjust dosages Adjust fluidized pellet bed Adjust by-pass ratio	Low	Low
Inactive activated carbon filters	Insufficient treatment due to low activity of activated carbon	Constant monitoring of contaminants and organic material passing through the carbon filter	Moderate	Low
Excessive regeneration of activated carbon	High environmental and financial costs	Constant monitoring of contaminants and organic material passing through the carbon filter	High	Moderate

The key treatment processes which are included in the water treatment module for every demo case vary and for DC#1 these are coagulation-flocculation, and ozonation. Each of the treatment steps represents a sub-module and aim to evaluate the ability of the system to consistently achieve water quality standards under any condition. The Water Treatment module will be connected with the Risk Assessment module in order to inform operators with alerts for the aforementioned key processes.

First, we considered coagulation-flocculation, a treatment technology intended to remove organic or particulate material. This is achieved by adding ferric chloride ($FeCl_3$), which forms large solid flocs: contaminants then adsorb to these flocs, which settle to the bottom of the tank allowing easy removal. Controlling the addition of $FeCl_3$ is important for both financial and environmental reasons: when too little is added, not enough removal of contaminants is achieved, but adding too much creates unnecessary use of chemicals. The turbidity at outlet is a parameter that operators use to optimize the $FeCl_3$ addition. A soft sensor is being developed that takes readily available data points and predicts turbidity at outlet, allowing for enhanced addition of $FeCl_3$ with better insight into the expected turbidity. With the expectation that $FeCl_3$ can be saved .

The second sub-module of the water treatment module for DC#1 is ozonation. Here, ozone is led through the water which reacts chemically with organic compounds and disinfects bacterial and viral matter. In this sub-module, the formation of assimilable organic carbon (AOC), the formation of bromate (BrO_3^-), and the disinfection capacity of ozonation are included. AOC functions as an indicator of the amount of organic matter in water that bacteria can assimilate and utilize for growth and metabolism. Monitoring of AOC concentration changes and notification, when AOC thresholds are exceeded, is important for operators. Additionally, during ozonation, it is important to monitor BrO_3^- formation, which occurs after the reaction of ozone with bromide. BrO_3^- is a disinfection by-product that is carcinogenic. Thus, monitoring the contact time of ozone with water, in combination with the disinfection capacity of ozonation to effectively inactivate pathogens, is essential and considered in the ozonation sub-module. The sub-module of ozonation can also receive data from the soft sensor at the end of ozonation process (developed in WP4 by TUD).

The key treatment processes included in the water treatment module for DC#2, where NTUA developed Reinforcement Learning (RL) models to optimize the processes of pre-chlorination, coagulation, and post-disinfection. The models use water quality parameters and follow a model-free approach that learns dosing policies that achieve target water quality variable thresholds with the minimum chemical input.

First, the pre-chlorination model provides the optimal chlorine dosage at the pre-disinfection stage, using as inputs the raw water turbidity at the inlet and the free chlorine concentration at the de-gritting stage and the filter's inlet. Second, the coagulation model provides the optimal aluminum sulphate dosage to maintain turbidity and free aluminum within safe operational limits, using the turbidity and temperature of the raw water, the turbidity at the filters' inlet, the free aluminum at the clear water tank inlet, and the free chlorine at the de-gritting stage. Third, the post-disinfection model proposes optimal chlorine dosage after filtration.

This way, the RL-derived dosing advice provides operators with direct operational recommendations indicating not only the expected water quality conditions but also the dosing actions required to maintain treatment effectiveness.

3.6 Distribution network

The distribution network module assesses risk of change in water quality related to leakage/burst, resulting in turbidity and/or increased concentration of pathogens, microbial regrowth – affecting concentrations of chlorine and associated DBPs. The aim here is to keep the water at a high drinking water quality, while avoiding the formation of DBPs (as much as possible).

In this module, Aeromonas concentration (See Table 17) will be included for DC#1 since it can be an indication of bacterial growth in the distribution system. Higher Aeromonas levels are connected with a higher organic nutrient source, which will lead to the growth of bacteria over time. The Aeromonas samples are measured at the outlet of the drinking water treatment plant. However, a daily prediction for Aeromonas is important for operators, especially in drinking water systems that do not use chlorine. Thus, a soft sensor that will predict the Aeromonas is aimed to be designed at the outlet of the drinking water treatment plant.

In **chlorinated systems**, chlorine residuals can deplete and reactions with organic matter can occur, leading to the formation of harmful DBPs.

In contrast, the main challenge in **non-chlorinated systems** is to maintain low AOC levels to prevent microbial regrowth, as they lack a disinfectant barrier. In these systems, the focus is on optimizing treatment processes to reduce organic matter as much as possible.

Each type of distribution system presents its own hazards and risks, which are outlined in detail in the Table 15 and 16 for chlorinated and non-chlorinated systems, respectively.

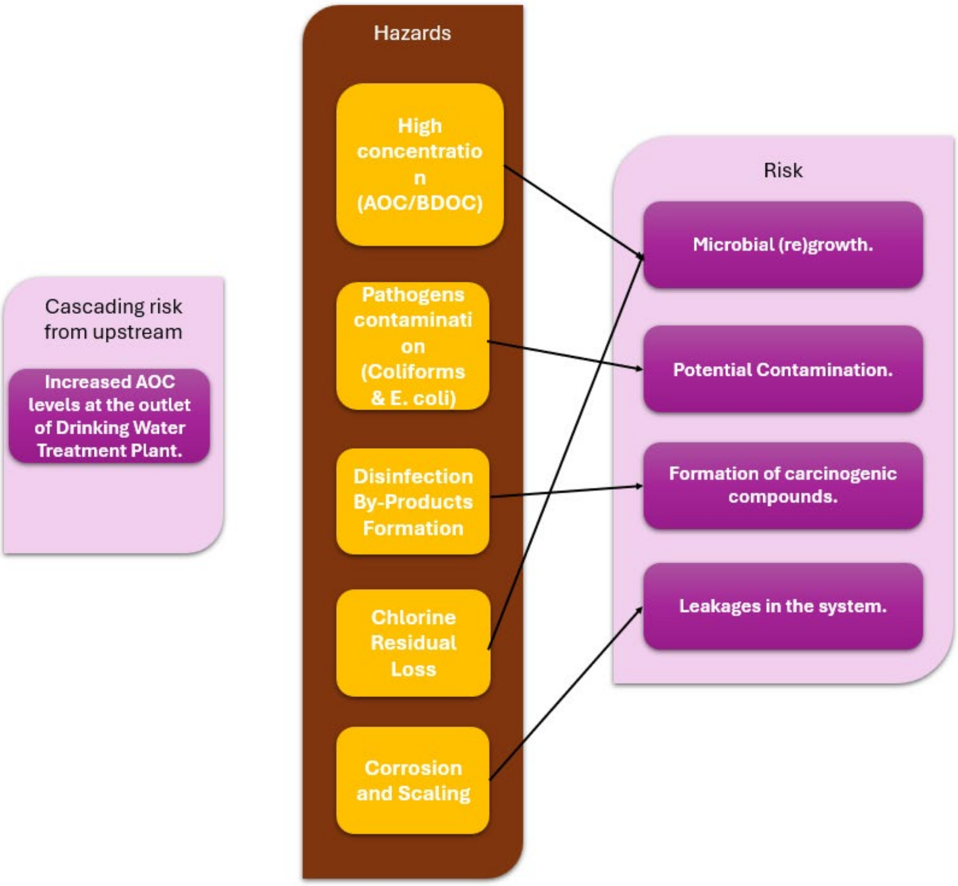


Figure 17: Mapping Chlorinated Distribution systems, potential hazards and potential risks

Table 15: Risk Identification for Chlorinated Distribution systems

Hazard	Potential Risk	Mitigation Strategy	Likelihood	Severity
Regrowth in the distribution network	AOC/BDOC concentration can lead to microbial regrowth in the distribution system, compromising water quality.	Monitor and control AOC/BDOC levels during WWTP processes to minimize concentration. Use advanced treatment if necessary.	Medium	High
Pathogens contamination (Coliforms & E. coli)	Presence of coliforms and E. coli indicates inadequate disinfection, leading to potential contamination.	Regularly monitor chlorine levels and ensure no coliforms or E. coli are present in the distribution system.	Low	High
Disinfection By-Products Formation	Formation of carcinogenic THMs and HAAs due to chlorine reacting with organic matter in the water.	Optimize chlorine dosage to minimize DBP formation. Regularly monitor THMs and HAAs levels.	High	High
Chlorine Residual Loss	Chlorine residual loss may allow microbial growth due to reaction with organic matter or depletion over time.	Maintain adequate chlorine residuals throughout the distribution system. Adjust chlorine dosing as necessary.	Medium	High

Hazard	Potential Risk	Mitigation Strategy	Likelihood	Severity
Corrosion and Scaling	Corrosion and scaling increase biofilm formation and lead to potential leakages in the system.	Regulate pH to prevent corrosion and scaling. Use corrosion inhibitors and maintain proper water chemistry.	Medium	High

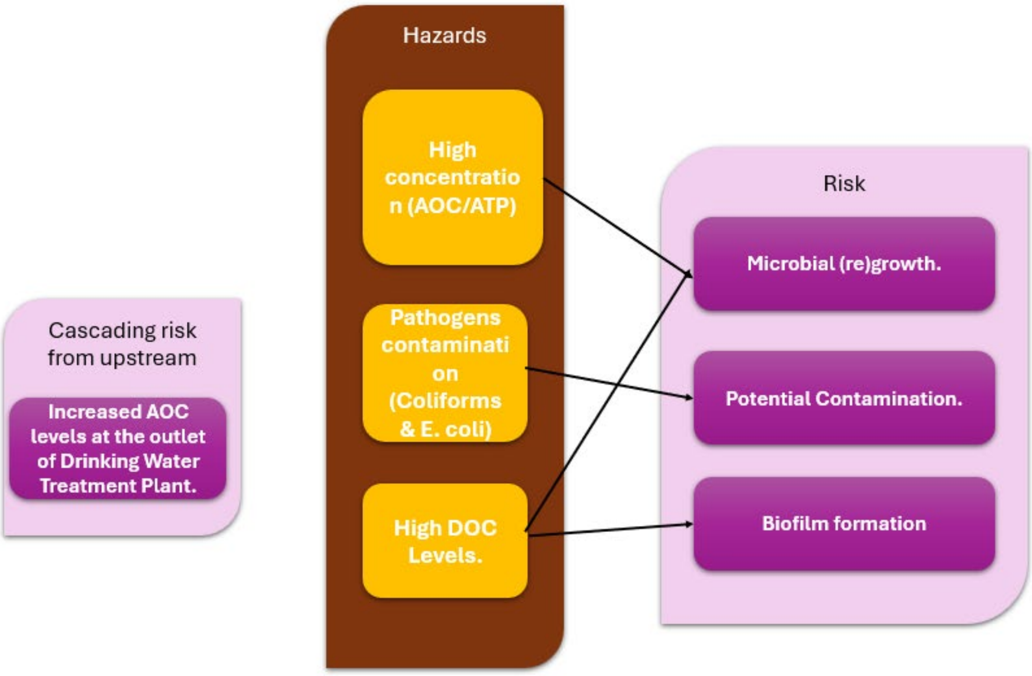


Figure 18: Mapping Non-Chlorinated Distribution systems, potential hazards and potential risks

Table 16: Risk Identification for Non-Chlorinated Distribution systems.

Hazard	Potential Risk	Mitigation Measures	Likelihood	Severity
Regrowth in the distribution network	Microbial regrowth due to AOC concentration, biofilm formation, and high AOC/ATP levels.	Monitor and control AOC levels, manage biofilms, and regularly monitor AOC/ATP levels. Use advanced treatment if necessary.	Moderate	High
	Microbial regrowth due to long residence times in the distribution system.	Optimize water residence time to reduce microbial growth and maintain flow.	Moderate	Moderate
	DOC promotes microbial growth and biofilm formation.	Reduce DOC through pre-treatment, filtration, or regular monitoring to minimize microbial growth.	Moderate	Moderate

Hazard	Potential Risk	Mitigation Measures	Likelihood	Severity
Pathogens contamination	Inadequate disinfection leading to contamination.	Regularly monitor water quality to ensure no E. coli or coliforms are present. Address leakage promptly.	Moderate	High
	Leakage or back flow leading to pathogen ingression	Regular maintenance, monitoring (e.g. pressure, visual), and inspection	Moderate	High
High DOC levels	Promotes microbial growth and biofilm formation.	Reduce DOC through pre-treatment, filtration, or regular monitoring to minimize microbial growth.	Moderate	Moderate

3.7 Sensors in the Modules

This section presents an overview of the sensors included in the ToDrinQ modules.

To support comprehensive monitoring and decision-making across the drinking water supply chain, the ToDrinQ modules integrate a set of soft sensors developed within WP4, complemented by hard sensors developed within WP3. The objective is to enable a holistic, data-driven assessment of water quality and operational performance, covering the full system from source to tap.

There is one generalized platform, NESSIE, which is the final platform deliverable of the ToDrinQ project (D7.4). This platform includes all modules, including hard sensors, soft sensors, and RL models presented in Table 17. NESSIE will include both demo cases. However, considering the specific operational needs of WTNT, a tailored FIWARE-compliant platform has also been developed. This platform is described in Chapter 3.8.

The current modules' implementation primarily focuses on the integration of soft sensors, which are fully operational within the FastAPI-based architecture and provide real-time predictions, risk evaluation, and decision support. These models are interconnected and operate sequentially across treatment stages, enabling a unified and dynamic representation of system behavior.

At the same time, hard sensors developed in ToDrinQ target key water quality parameters (e.g., pathogens, nutrients, and heavy metals) and are expected to further enhance monitoring capabilities once fully validated and deployed. However, only BactoSense (hard sensor #4) could be uploaded on the platform because it can provide real time data. It is further disclaimed on the D.7.3, after your useful comment, that since WTNT prefers having a "validation" and "testing" period of this hard sensor before the hosting on their PIMS (Process Information Management System). For this reason, currently the BactoSense will be hosted by ATP (a platform provider). For DC#2 (EYDAP), discussions are still ongoing regarding which hard sensors (existing or the ones developed within ToDrinQ) should be displayed based on operational requirements.

For the ToDrinQ modules of the platform, there are two possible scenarios concerning the inclusion of the hard sensors: (1) we connect with ATP platform with an API to ToDrinQ platform or (2) take a csv of historical data creates fake data (for respecting the confidentiality) of data output of the sensors based on real historical data, and displaying the way that the BactoSense (hard sensor #4) could possibly be integrated in the future on the platform after the "validation period" which there would not be connection with the PIMS.

To ensure clarity and transparency, Table 17 provides a detailed overview of all sensors considered within the modules, including their classification (soft/hard), work package origin, and implementation status.

This overview is presented in a structured table, where all sensors are consistently named in alignment with the Description of Action (DoA).

The table reflects the **final selection and alignment with end-user requirements (WTNT and EYDAP)**, as well as the implementation status within the platform. It provides a structured overview of all sensors, including:

- their classification (**soft or hard sensor**),
- their **work package of development (WP3/WP4)**,
- their mapping to the **corresponding module (source to tap)**,
- their **functional role** (prediction or monitoring),
- and their **integration status within the platform**.

Table 17: Sensors and RL models on the modules of the platform DC #1 WTNT and DC #2 EYDAP.

Sensor Name	***Type***	***Work Package***	***Module (from source to tap)***	***Description***	***Status in Platform***	***Water Utility***
Turbidity inlet prediction model	*Soft sensor* ***SoSe #9***	*WP4*	*Source / Intake*	*Predicts raw water turbidity up to 3 hours ahead*	*Implemented* *In the platform*	*WTNT* ***DC#1***
Coagulation performance model	*Soft sensor* ***Sose #8***	*WP4*	*Treatment - Coagulation*	*Predicts turbidity after coagulation*	*Implemented* *In the platform*	*WTNT* ***DC#1***
RL coagulant optimization model	*Reinforcement Learning* ***RL #5***	*WP4*	*Treatment - Coagulation*	*Recommends optimal $FeCl_3$ dosage*	*Implemented* *In the platform*	*WTNT* ***DC#1***
CT (ozone exposure) model	*Soft sensor* ***SoSe #10***	*WP4*	*Treatment - Ozonation*	*Estimates disinfection capacity (CT)*	*Implemented* *In the platform*	*WTNT* ***DC#1***
Bromate formation model	*Soft sensor* ***SoSe #10.1***	*WP4*	*Treatment - Ozonation*	*Predicts bromate formation risk*	*Implemented* *In the platform*	*WTNT* ***DC#1***
Coliform removal model	*Soft sensor* ***SoSe #10.2***	*WP4*	*Treatment - Ozonation*	*Predicts microbial removal efficiency*	*Implemented* *In the platform*	*WTNT* ***DC#1***
Aeromonas prediction model	*Soft sensor* ***SoSe #11***	*WP4*	*Distribution*	*Predicts the likelihood and dynamics of Aeromonas occurrence in the distribution system, acting as an indicator of biological*	*Under evaluation / subject to change*	*WTNT* ***DC#1***

Sensor Name	***Type***	***Work Package***	***Module (from source to tap)***	***Description***	***Status in Platform***	***Water Utility***
				instability and potential microbial regrowth in treated drinking water.		
Bactosense	*Hard sensor* ***#4***	*WP3*	*Treatment*		*Implemented* *In the platform*	*WTNT* ***DC#1***
Early Warning for nutrient runoff in the Yliki Lake	*Soft sensor* ***SoSe #1***	*WP4*	*Source*	It provides an early warning system for nutrient runoff for in Yliki lake.	*Implemented* *In the platform*	*EYDAP* ***DC#2***
Chlorophyl-a estimation	*Soft sensor* ***SoSe #2***	*WP4*	*Source*	It provides Chl-a concentration estimations in frequent time intervals (3-5 days) at the pixel level (10m).	*Implemented* *In the platform*	*EYDAP* ***DC#2***
pH estimation	*Soft sensor* ***SoSe #3***	*WP4*	*Source*	It provides pH estimations in frequent time intervals (3-5 days) at the pixel level (10m)	*Implemented* *In the platform*	*EYDAP* ***DC#2***
DO estimation	*Soft sensor* ***SoSe #4***	*WP4*	*Source*	It provides DO concentration estimations in frequent time intervals (3-5 days) at the pixel level (10m)	*Implemented* *In the platform*	*EYDAP* ***DC#2***
Bloom Occurrence Probability	*Soft sensor* ***SoSe #5***	*WP4*	*Source*	It estimates of the Floating Algal bloom occurrence probability at pixel level	*Implemented* *In the platform*	*EYDAP* ***DC#2***
Water Quality Index (WQI)	*Soft sensor* ***SoSe #6***	*WP4*	*Source*	It estimates the Water Quality Index (WQI) in the Yliki lake at pixel level	*Implemented* *In the platform*	*EYDAP* ***DC#2***
Optimization of pre-disinfection stage	*Reinforcement Learning* ***RL #1***	*WP4*	*Treatment*	The RL-module recommends disinfectant dosage for the pre-disinfection treatment process and domain experts can validate the RL dosage recommendation.	*Implemented* *In the platform*	*EYDAP* ***DC#2***
Optimization of	*Reinforcement Learning*	*WP4*	*Treatment*	The RL-module recommends Coagulant	*Implemented*	*EYDAP*

Sensor Name	***Type***	***Work Package***	***Module (from source to tap)***	***Description***	***Status in Platform***	***Water Utility***
coagulation stage	***RL #2***			dosage for the coagulation/Flocculation treatment process and domain experts can validate the RL dosage recommendation.	*In the platform*	***DC#2***
Optimization of post-disinfection stage	*Reinforcement Learning* ***RL #3***	*WP4*	*Treatment*	The RL-module recommends disinfectant dosage for the post-disinfection treatment process and domain experts can validate the RL dosage recommendation.	*Implemented* *In the platform*	*EYDAP* ***DC#2***

For further clarification, the following tables (Table 18 and 19) provide an overview of the hard sensors, soft sensors, and RL models associated with each module of the water supply chain and their connection to risk assessment for both demo cases (DC#1 and DC#2). These tables are intended to complement the information presented in Table 18.

Table 18 Mapping of DC#1 (WTNT) Sensors and Models to Platform Modules and Their Contribution to the Risk Assessment Module.

Module	**Sensors / Models**	**Risk Assessment Input**
Source / Intake	Turbidity Inlet Prediction Model (SoSe #9). It should also be noted that the tailored WTNT platform includes a live connection to the Rijkswaterstaat Waterinfo service, from which real-time data are retrieved for stations relevant to WTNT (e.g., Lekkanaal). However, this connection serves only as an external data source and does not represent a hard sensor developed within the ToDrinQ project. This demonstrates the platform's capability to integrate external IoT-enabled and online data sources alongside ToDrinQ-developed sensors.	Raw water turbidity risk
Conveyance	The conveyance module, which represents the transfer of water between the source and the treatment facility, is part of the overall NESSIE architecture. However, based on discussions with WTNT stakeholders, no operational requirement was identified for its inclusion in the tailored WTNT (DC#1) platform. Consequently, no conveyance-related sensors or risk indicators are displayed in the WTNT implementation. Nevertheless, the conveyance module remains available in the final NESSIE platform and can support demonstration case DC#2 (EYDAP).	-
Treatment	Coagulation Performance Model (SoSe #8), RL Coagulant Optimization Model (RL #5)	Risk of insufficient turbidity removal

Module	Sensors / Models	Risk Assessment Input
Treatment	CT Model (SoSe #10), Bromate Formation Model (SoSe #10.1), Coliform Removal Model (SoSe #10.2)	Disinfection failure risk, bromate formation risk, microbial removal risk
Treatment	BactoSense Hard Sensor (#4)	Microbiological contamination risk
Distribution System	Aeromonas Prediction Model (SoSe #11) (under evaluation and subject to change)	Biological instability and microbial regrowth risk

Table 19: Mapping of DC#2 (EYDAP) Sensors and Models to Platform Modules and Their Contribution to the Risk Assessment Module.

Module	Sensors / Models	Risk Assessment Input
Source / Intake	Early Warning System for nutrient runoff (SoSe #1)	Risk of nutrient runoff based on runoff and nutrient loading on the upstream basin
Source / Intake	Chl-a concentration estimation (SoSe #2)	Risk of eutrophication based on chlorophyll-a, complicating treatment process
Source / Intake	pH estimation (SoSe #3)	Risk of reduced disinfection efficiency from pH outside the optimal range
Source / Intake	DO concentration estimation (SoSe #4)	Risk of anoxic conditions from low dissolved oxygen, promoting odor and taste problems in treatment.
Source / Intake	Bloom Occurrence Probability (SoSe #5)	-
Source / Intake	Water Quality Index (WQI) (SoSe #6(	-
Conveyance	Installed Turbidity sensor	Risk of high suspended-solids and sediment loading from elevated turbidity, overloading coagulation and filtration.
Conveyance	Installed Conductivity sensor	Risk of elevated salinity or dissolved solids accelerating pipe corrosion and complicating treatment process
Treatment	Optimization of pre-disinfection stage (RL #1)	-
Treatment	Optimization of coagulation stage (RL #2)	-
Treatment	Optimization of post-disinfection stage (RL #3)	-
Distribution System	It is important to note that, for EYDAP (DC#2), no soft sensor or other monitoring tool has currently been identified for the distribution module. Nevertheless, the platform	-

Module	Sensors / Models	Risk Assessment Input
	architecture fully supports this module. For WTNT (DC#1), the distribution module is currently represented by the Aeromonas Prediction Model (SoSe #11), which is under evaluation and may be subject to future modifications.	

3.8 *Modules architecture (backend) DC#1*

The ToDrinQ platform constitutes the software environment through which the developed prediction models, monitoring information, and risk assessment functionalities are operationalized into a unified decision-support system for drinking water utilities. While the previous sections describe the individual modules from a conceptual and functional perspective, the present section explains how these modules are implemented within the backend of the platform and how they are interconnected in practice.

As already mentioned, the platform development has been guided by the end-user requirements, with Delft University of Technology (TUD) following the specific needs of WTNT. Since WTNT already operates its own legacy system, the direct adoption of NESSIE was not considered practical for their operational environment.

As a result, the proposed solution was the development of a FIWARE-aligned Application Programming Inteface (API)-based platform capable of:

1. Connecting with NESSIE for project-level integration and demonstration purposes,
2. Interfacing with the existing WTNT legacy system,
3. Incorporating the modules developed for DC# 1 (WTNT), namely the **Source**, **Treatment**, **Distribution**, and **Risk** modules, and
4. Be able to be exploited in the future for other demo cases (other water utilities).

Within the NESSIE environment, both demonstration cases (DC# 1: WTNT and DC# 2: EYDAP) will be represented. However, the focus of this chapter is on the platform developed by TUD for the WTNT use case.

The platform has been implemented as a Dockerized FastAPI-based backend, designed to support real-time or near real-time monitoring, predictive analysis, and risk-based operational support. Its main objective is to transform raw operational data into actionable information for operators, enabling the transition from reactive monitoring to proactive system management. The backend hosts the prediction models, the preprocessing logic, the risk assessment module, and the API endpoints that facilitate communication with external systems.

For simplicity, this Dockerized FastAPI-based backend is hereafter referred to as the **ToDrinQ platform**.

For further information, you can send us email and send you an invitation. Please contact Akrivi Alexandraki or Dr. Greg Kyritsakas to the following email addresses: A.A.Alexandraki@tudelft.nl or G.Kyritsakas@tudelft.nl. On the final publication of the platform the repository will be uploaded on Zenodo. For now, we protect our models by keeping our modules private. Because of the random values as input, maybe some of the results will not be logical in the following examples given in screenshots from the platform.

3.8.1 Backend architecture

The platform follows a modular and scalable architecture aligned with FIWARE principles. At a high level, the backend is organized into the following layers:

1. **Data source layer:**
Operational data are collected from the utility infrastructure, primarily through SCADA systems and the Process Information Management System (PIMS). These data include conventional parameters such as turbidity, flow, temperature, pH, ozone dosage, and other operational settings, and are used as inputs for the soft sensors hosted on the platform. In addition, this layer provides the basis for integrating hard sensor measurements once they become available and are connected to the operational environment.
2. **Connector layer:**
A connector retrieves the required input variables from PIMS and transfers them to the platform through HTTP/API requests. In this way, the connector acts as the interface between the legacy water utility systems and the ToDrinQ platform.
3. **Core platform layer:**
The core of the platform is implemented in FastAPI. This application hosts:
    - the soft sensor prediction models,
    - the preprocessing and feature engineering logic,
    - the risk assessment logic,
    - the REST API endpoints used for model execution and communication.

   Each model is exposed as a dedicated API endpoint, enabling flexible integration and modular deployment.

4. **Output and feedback layer:**

The backend returns structured JSON responses containing prediction values, risk indicators, alarm flags, alarm descriptions, and suggested mitigation actions. These outputs can be displayed in the user interface and, in the final operational setup, can also be returned through the connector to the utility environment for operational use.

The overall operational workflow of the platform is therefore:

SCADA → PIMS → Connector → FastAPI platform → Soft sensor models → Risk assessment → Output → User Interface / PIMS

This structure ensures that the modules are not only described conceptually, but are also implemented in software through a sequence of connected processing steps.

In addition to the layered architecture, the platform is structured into four functional modules: **Source**, **Treatment**, **Distribution**, and **Risk**, which operate within the core platform layer. These modules host the corresponding soft sensors and interact through the defined API endpoints, ensuring interoperability and alignment with the source-to-tap logic of drinking water systems.

To provide a comprehensive representation of the backend architecture and its integration within the operational environment, Figure 19 presents the overall system design of the ToDrinQ platform. The figure illustrates how the different architectural layers: data source, connector, core platform, and output, are implemented and interact in practice.

In addition, the figure highlights the deployment of the platform as a Dockerized FastAPI application, as well as its interoperability with external systems, including the WTNT legacy infrastructure and the NESSIE environment. The internal structure of the platform is also depicted, showing the implementation of the **Source, Treatment, and Distribution modules as prediction services**, and the **Risk module as an aggregation and decision-support layer**.

Overall, the figure complements the textual description by demonstrating the end-to-end data flow, from data acquisition (SCADA/PIMS) to prediction, risk assessment, and feedback to the user interface and operational systems, thereby confirming the implementation of a fully functional, modular, and interoperable backend system.

In the following section, the organization of the modules and their integration with soft and hard sensors is further explained.

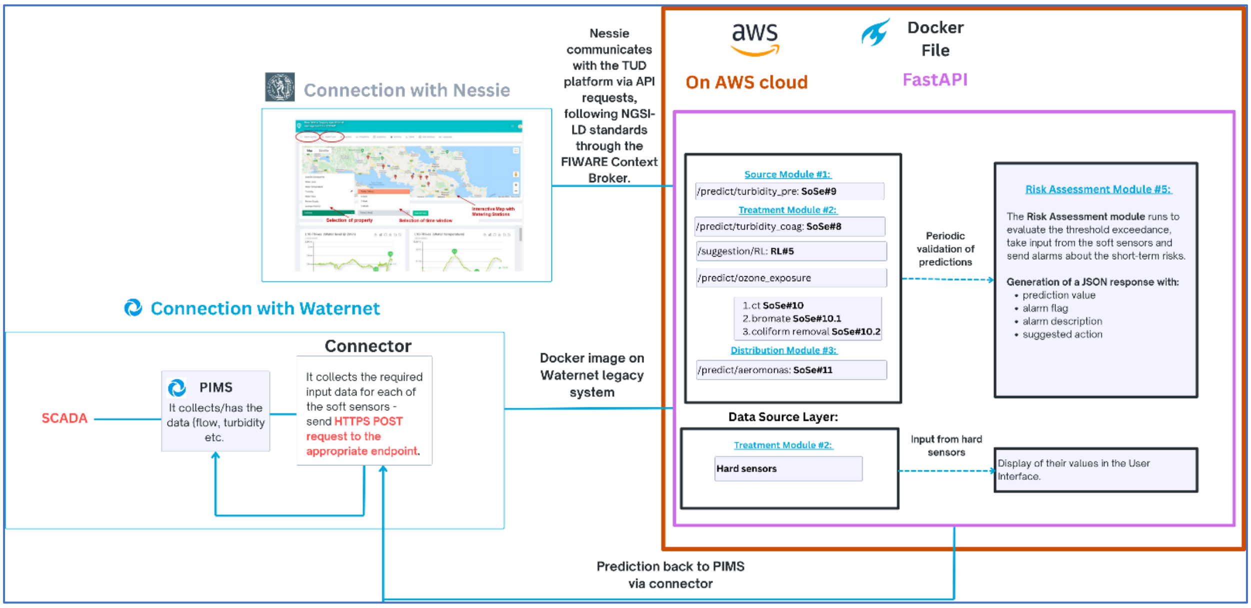


Figure 19: Backend architecture of the ToDrinQ platform, showing data flow from SCADA/PIMS through the connector to the Dockerized FastAPI platform, where Source, Treatment, Distribution, and Risk modules are implemented

The figure also illustrates interoperability with WTNT and NESSIE environments.

3.8.2 Modules from Source to Tap

For DC# 1 (WTNT), four main modules are implemented: **Source**, **Treatment**, **Distribution**, and **Risk**. Each module incorporates the corresponding soft sensors and, where applicable, hard sensors, aligned with the specific purpose and functionality of the module.

For example, the soft sensor related to **Concentration Time (CT) exposure** (SoSe #10) is implemented within the **Treatment module**. This soft sensor consists of a set of interconnected prediction models and is designed to evaluate the efficiency of the ozonation process, which is a key disinfection method used in the Netherlands.

Since this sensor directly assesses treatment performance, specifically the effectiveness of disinfection, it is logically integrated within the Treatment module. This illustrates how sensors are allocated to modules based on their functional role within the overall drinking water treatment process.

To further demonstrate the implementation of the modular architecture within the backend, the following Figures present the FastAPI Swagger interface of the ToDrinQ platform. These screenshots provide direct evidence of the developed modules as executable and accessible API services.

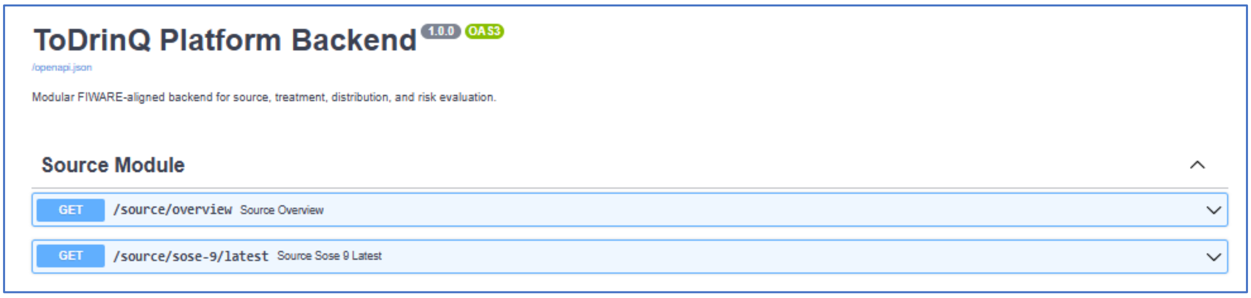
Each module, **Source**, **Treatment**, **Distribution**, and **Risk**, is represented as a dedicated group of endpoints, confirming that the modular structure is not only conceptual but fully implemented in the backend. The endpoints are organized according to their functional role and follow a consistent naming structure (e.g., /source/..., /treatment/..., /distribution/..., /risk/...), enabling clear separation of responsibilities and interoperability between modules.

The **Source module** includes endpoints related to raw water conditions and upstream predictions. The **Treatment module** contains multiple endpoints corresponding to the different stages of the treatment process, including coagulation and ozonation models (e.g., CT exposure, bromate formation, and coliform removal), as well as integration points for hard sensors. The **Distribution module** hosts endpoints associated with water quality in the distribution network.

Finally, the **Risk module** aggregates the outputs of all other modules and provides endpoints for evaluating risk at different levels, including individual sensors, treatment-level assessment, and full **source-to-tap risk evaluation**. This module demonstrates how predictions are transformed into actionable decision-support outputs.

Overall, these figures confirm that the platform has been implemented as a modular, API-based system, where each module operates as an independent yet interconnected service. This design ensures interoperability, scalability, and seamless integration with external systems such as PIMS, WTNT's legacy infrastructure, and the NESSIE environment.

These figures collectively demonstrate that the ToDrinQ platform modules are fully implemented as interoperable API services, fulfilling the WP7 requirement for modular, scalable, and integrable digital components.



ToDrinQ Platform Backend 1.0.0 OAS3

/openapi.json

Modular FIWARE-aligned backend for source, treatment, distribution, and risk evaluation.

Source Module

GET `/source/overview` Source Overview

GET `/source/sose-9/latest` Source Sose 9 Latest

Figure 20: FastAPI Swagger interface showing the implemented endpoints of the ***Source module***

The module includes routes for retrieving source-level information and accessing the latest predictions from the inlet turbidity soft sensor (SoSe #9). This demonstrates the implementation of the source component as an independent, API-accessible service within the platform.

Treatment Module

GET `/treatment/overview` Treatment module overview

GET `/treatment/coagulation/latest` Treatment SoSe #8 latest

GET `/treatment/coagulation/rl/latest` Treatment RL #5 latest

POST `/treatment/ozonation/ct` Treatment SoSe #10 CT

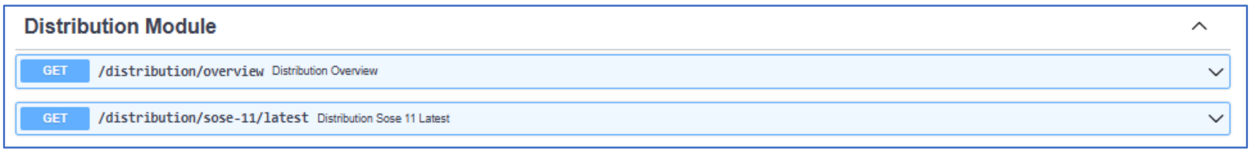
POST `/treatment/ozonation/bromate` Treatment SoSe #10.1 bromate

POST `/treatment/ozonation/coliform` Treatment SoSe #10.2 coliform

GET `/treatment/hard-sensors/bactosense-4/latest` Treatment BactoSense #4 latest

*Figure 21: FastAPI Swagger interface showing the **Treatment module** endpoints*

The module integrates multiple soft sensors, including coagulation performance, reinforcement learning for coagulant dosing, and ozonation-related models (CT exposure, bromate formation, and coliform removal), as well as hard sensor integration (e.g., BactoSense). The figure highlights the complexity and interconnection of prediction services within the treatment process.



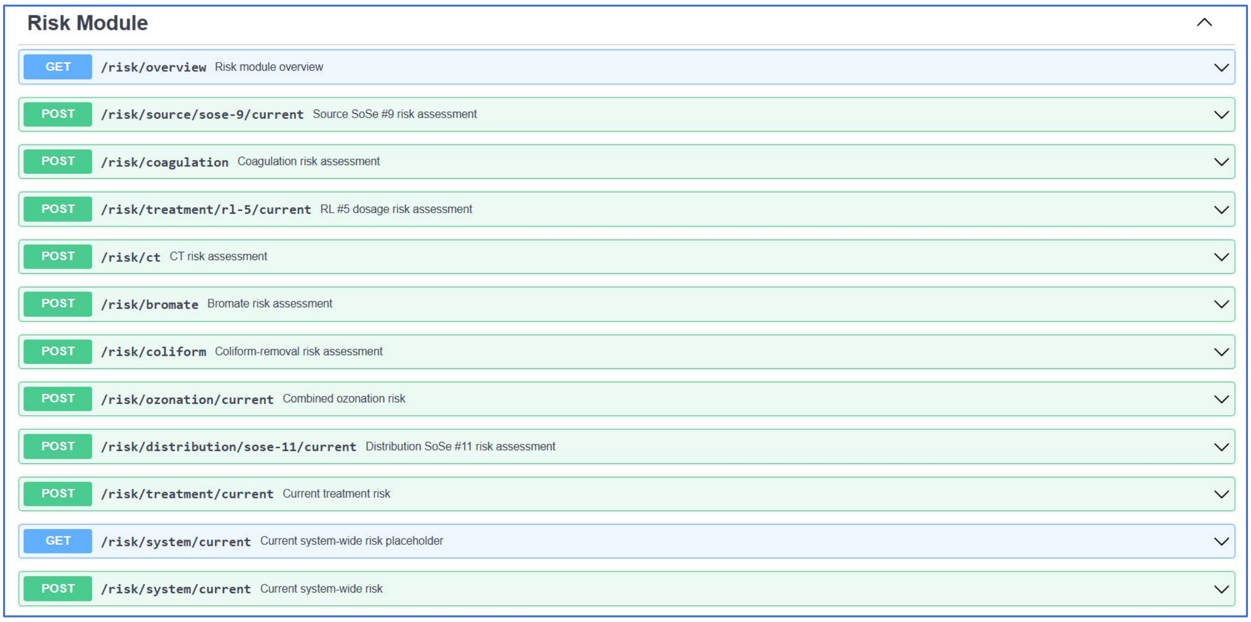
Distribution Module

GET `/distribution/overview` Distribution Overview

GET `/distribution/sose-11/latest` Distribution Sose 11 Latest

*Figure 22: FastAPI Swagger interface presenting the **Distribution module** endpoint*

The module includes routes for monitoring and prediction of water quality within the distribution network, such as the Aeromonas-related soft sensor (SoSe #11). This confirms the implementation of distribution-level monitoring and predictive capabilities within the platform.



Risk Module

GET `/risk/overview` Risk module overview

POST `/risk/source/sose-9/current` Source SoSe #9 risk assessment

POST `/risk/coagulation` Coagulation risk assessment

POST `/risk/treatment/rl-5/current` RL #5 dosage risk assessment

POST `/risk/ct` CT risk assessment

POST `/risk/bromate` Bromate risk assessment

POST `/risk/coliform` Coliform-removal risk assessment

POST `/risk/ozonation/current` Combined ozonation risk

POST `/risk/distribution/sose-11/current` Distribution SoSe #11 risk assessment

POST `/risk/treatment/current` Current treatment risk

GET `/risk/system/current` Current system-wide risk placeholder

POST `/risk/system/current` Current system-wide risk

*Figure 23: FastAPI Swagger interface showing the **Risk module**, which aggregates outputs from all other modules*

The endpoints include risk evaluation for individual sensors, treatment-level risk, and full **source-to-tap risk assessment**. This figure demonstrates how the platform translates predictions into risk indicators, alarms, and decision-support outputs.

3.8.3 Integration of the soft sensors in the modules

All soft sensor models described in the previous sections are integrated into the backend and exposed as API endpoints. These models operate within a unified framework and can be executed sequentially where required. This is particularly important for the treatment-stage models, where the output of one model becomes input to the next one.

At the current stage, the following soft sensors are integrated or under active evaluation in the platform:

- turbidity inlet prediction model,
- coagulation performance model,
- reinforcement learning coagulant optimization model,
- ozone exposure (CT) model,
- bromate formation model,
- coliform removal model,
- Aeromonas prediction model (under evaluation / subject to change).

The backend enables the sequential and interconnected execution of these models across the treatment process. This is especially relevant in ozonation, where the CT model must run first and its result is then used by the bromate formation model and the coliform removal model. In this way, the platform does not simply host isolated models, but supports an integrated decision-support logic across stages of the treatment process .

3.8.4 Full pipeline of all models

On the platform, both hard and soft sensors developed within ToDrinQ are hosted, while additional capacity is provided to integrate new sensors in the future. This ensures that the platform is designed as a scalable and sustainable solution.

Furthermore, the requirement from WTNT to develop a platform that can be integrated into their existing legacy system initially posed a challenge; however, it also represents a significant opportunity for the future exploitation of the platform across additional demonstration cases and other water utilities.

In this section, the pipeline of all models currently developed within ToDrinQ, presented in Table 17 of Deliverable D7.2, is analyzed. The pipeline begins with raw water conditions and follows the operational logic of the treatment process.

The implemented pipeline follows the operational logic of the drinking water system and is structured according to the **Source, Treatment, Distribution, and Risk modules** of the platform. The models are executed sequentially across these modules, enabling an integrated source-to-tap decision-support workflow.

RAW WATER (Inlet Conditions) → SOURCE MODULE

├Turbidity

├Temperature

├pH

├Flow

▼

COAGULATION STAGE → TREATMENT MODULE

Coagulation Performance Model

→ Predicts turbidity after coagulation

RL Coagulant Optimization Model

→ Suggests optimal $FeCl_3$ dosage

OUTPUT:

✓ Predicted effluent turbidity

✓ Recommended coagulant dose

▼

OZONATION STAGE → TREATMENT MODULE

Inputs:

\- Flow

\- Turbidity (from coagulation / inlet)

\- Temperature

\- Ozone dosage

\- Chamber

Figure 24: Full pipeline of the models on the platform and their connection is given (Part 1)

▼
Ozone Exposure Model (CT)
→ Calculates CT (disinfection capacity)

OUTPUT:
✓ CT ozone exposure
▼
Bromate Formation Model
→ Predicts bromate concentration

Inputs:
- CT exposure (from previous model)
- + time-lagged variables
- + operational conditions

OUTPUT:
✓ Bromate concentration

▼
Microbial Removal Model
→ Predicts log removal of coliforms

Inputs:
- CT exposure
- Turbidity
- Flow
- Temperature

OUTPUT:
✓ Log removal of pathogens

FINAL PLATFORM OUTPUT → RISK MODULE

✓ Water quality indicators
✓ Disinfection performance
✓ By-product risk (bromate)
✓ Recommended actions (dose optimization)
✓ Alerts & decision support

Figure 25: Full pipeline of the models on the platform and their connection is given (Continuation from Part 1 – Part 2)

3.8.4.1 Raw water / inlet conditions (Source module)

The initial input space includes conventional operational variables such as turbidity, temperature, pH, and flow, which are retrieved from the data source layer (SCADA/PIMS) through the connector. Within the **Source module**, the turbidity inlet prediction model estimates raw water turbidity ahead of time, enabling the anticipation of changes in raw water quality.

The output of this module is exposed through an API endpoint and serves as input for downstream processes within the Treatment module.

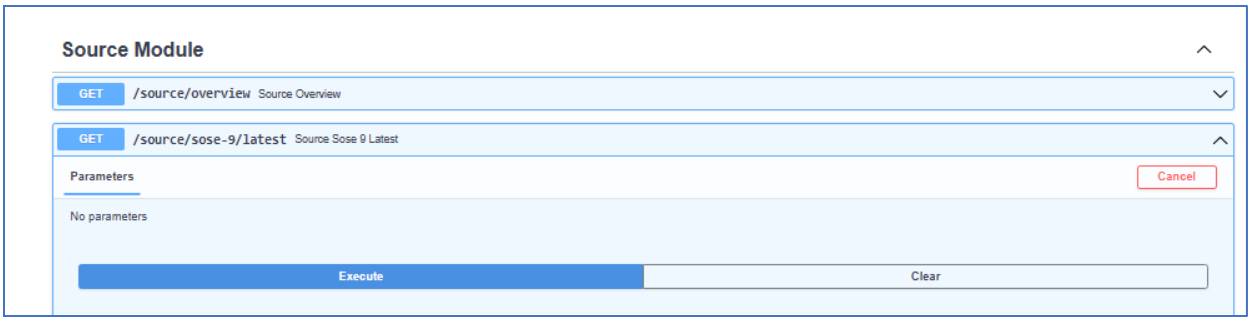
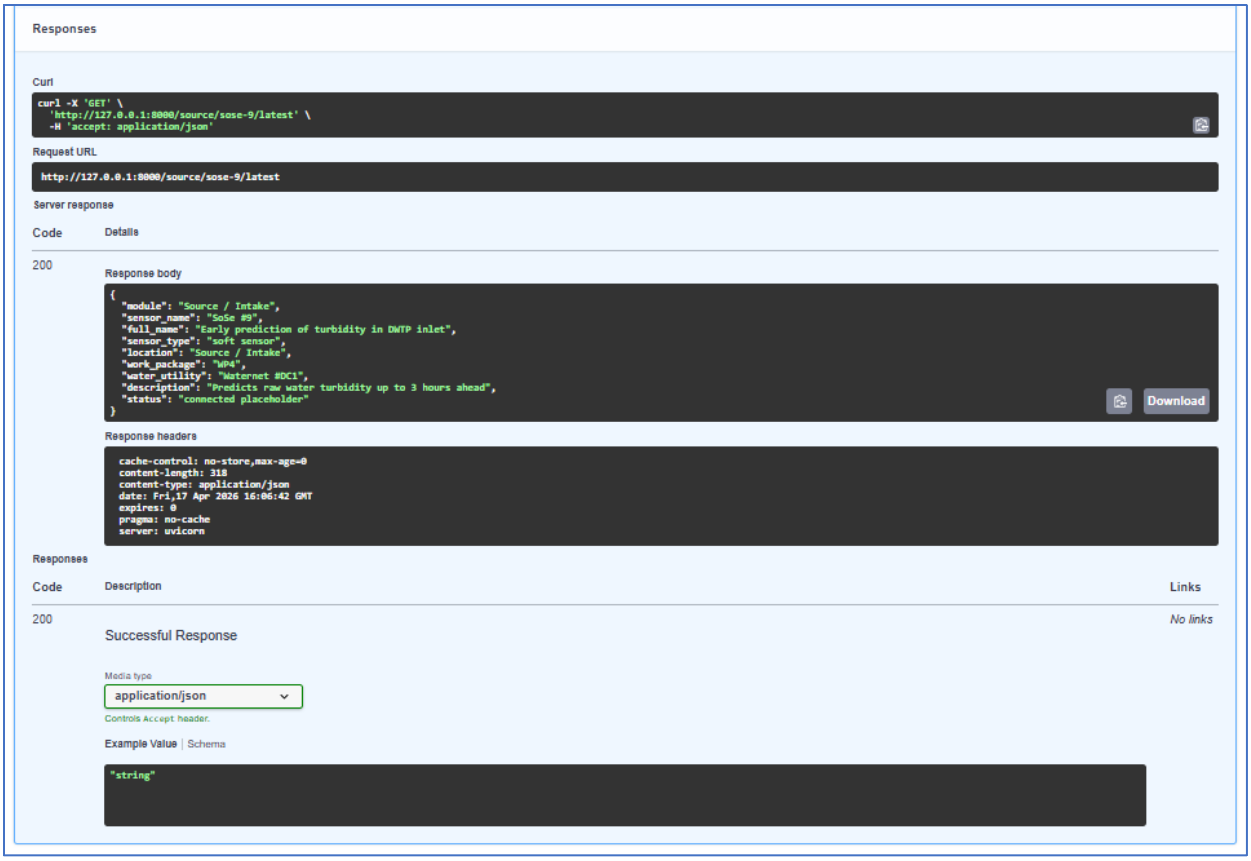


Figure 26: FastAPI Swagger interface showing the Source module endpoints, including the source overview and the latest output of the turbidity inlet soft sensor (SoSe #9)

This demonstrates the implementation of the Source module as an API-accessible component within the platform.

If the user clicks on execute, the subsequent figure shows the response body of the source module which includes the SoSe #9.



```
curl -X 'GET' \
  'http://127.0.0.1:8000/source/sose-9/latest' \
  -H 'accept: application/json'
```

```
http://127.0.0.1:8000/source/sose-9/latest
```

```
{
  "module": "Source / Intake",
  "sensor_name": "SoSe #9",
  "full_name": "Early prediction of turbidity in DWTP inlet",
  "sensor_type": "soft sensor",
  "location": "Source / Intake",
  "work_package": "WP4",
  "water_utility": "Waternet #DC1",
  "description": "Predicts raw water turbidity up to 3 hours ahead",
  "status": "connected placeholder"
}
```

```
 cache-control: no-store,max-age=0
 content-length: 318
 content-type: application/json
 date: Fri,17 Apr 2026 16:06:42 GMT
 expires: 0
 pragma: no-cache
 server: uvicorn
```

```
"string"
```

Figure 27: Example execution of the Source module endpoint in the FastAPI backend, showing the retrieval of the latest output from the turbidity inlet soft sensor (SoSe #9)

The structured JSON response confirms the implementation and accessibility of the Source module through API-based communication.

3.8.4.2 Coagulation stage (Treatment module)

Within the **Treatment module**, the coagulation performance model predicts the expected effluent turbidity after coagulation, while the reinforcement learning model recommends an optimized ferric chloride dosage.

The outputs of this stage are:

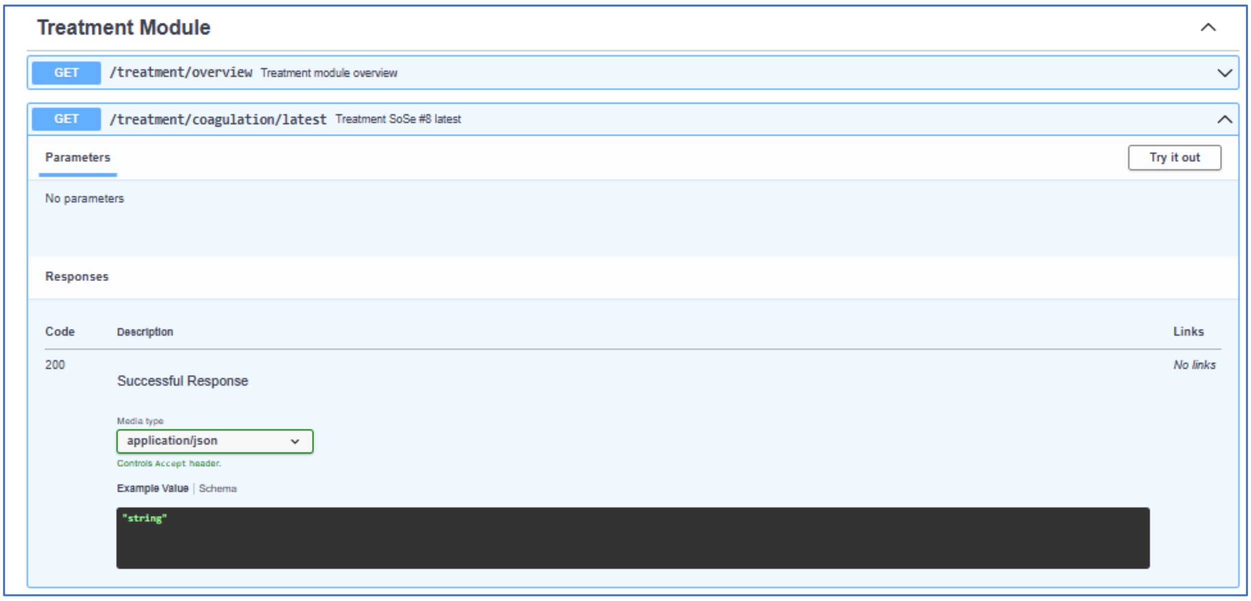
- predicted turbidity after coagulation,
- recommended coagulant dosage.

These outputs are made available through dedicated API endpoints and are used as inputs for subsequent treatment-stage models.

Treatment Module

GET `/treatment/overview` Treatment module overview
GET `/treatment/coagulation/latest` Treatment SoSe #8 latest
GET `/treatment/coagulation/rl/latest` Treatment RL #5 latest
POST `/treatment/ozonation/ct` Treatment SoSe #10 CT
POST `/treatment/ozonation/bromate` Treatment SoSe #10.1 bromate
POST `/treatment/ozonation/coliform` Treatment SoSe #10.2 coliform
GET `/treatment/hard-sensors/bactosense-4/latest` Treatment BactoSense #4 latest

Figure 28 Treatment module which includes the Soft Sensors SoSe #8, 10, 10.1, 10.2, RL #5 and Hard sensor: BactoSense #4



Treatment Module

GET `/treatment/overview` Treatment module overview
GET `/treatment/coagulation/latest` Treatment SoSe #8 latest

Parameters

No parameters

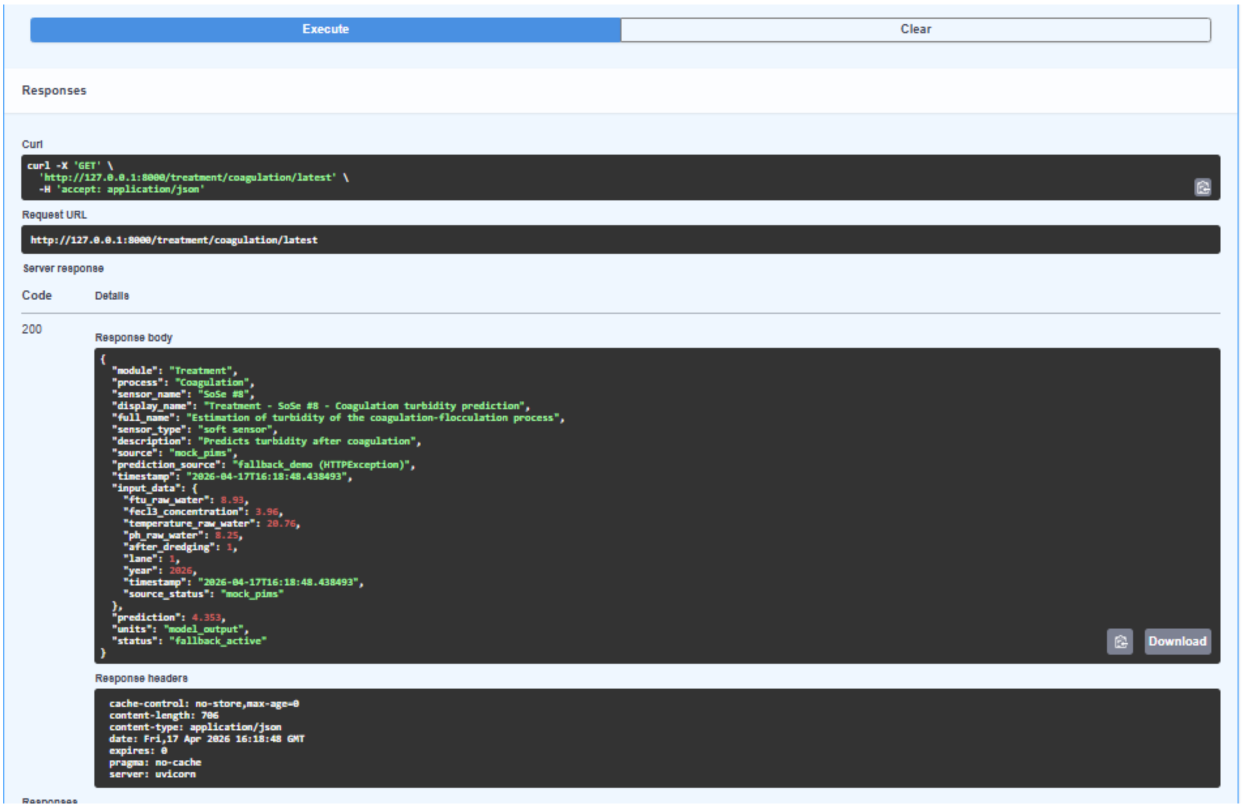
Responses

Code	Description	Links
200	Successful Response Media type: application/json Controls Accept header. Example Value | Schema "string"	No links

Figure 29: FastAPI Swagger interface showing the Treatment module endpoints, including the coagulation prediction service

The figure demonstrates the availability of treatment-stage functionalities through API endpoints within the platform.

If the user clicks on the Try it out, the subsequent figure shows the response body for the SoSe #8.



```
curl -X 'GET' \
  'http://127.0.0.1:8000/treatment/coagulation/latest' \
  -H 'accept: application/json'
```

```
http://127.0.0.1:8000/treatment/coagulation/latest
```

```
{
  "module": "Treatment",
  "process": "Coagulation",
  "sensor_name": "SoSe #8",
  "display_name": "Treatment - SoSe #8 - Coagulation turbidity prediction",
  "full_name": "Estimation of turbidity of the coagulation-flocculation process",
  "sensor_type": "soft sensor",
  "description": "Predicts turbidity after coagulation",
  "source": "mock_pims",
  "prediction_source": "fallback_demo (HTTPException)",
  "timestamp": "2026-04-17T16:18:48.438493",
  "input_data": {
    "ftu_raw_water": 8.93,
    "fecl3_concentration": 3.96,
    "temperature_raw_water": 20.76,
    "ph_raw_water": 8.25,
    "after_dredging": 1,
    "lane": 1,
    "year": 2026,
    "timestamp": "2026-04-17T16:18:48.438493",
    "source_status": "mock_pims"
  },
  "prediction": 4.353,
  "units": "model_output",
  "status": "fallback_active"
}
```

```
 cache-control: no-store,max-age=0
 content-length: 706
 content-type: application/json
 date: Fri,17 Apr 2026 16:18:48 GMT
 expires: 0
 pragma: no-cache
 server: uvicorn
```

Figure 30: Example execution of the coagulation prediction endpoint within the Treatment module, showing input parameters retrieved from the data source layer and the corresponding structured JSON output

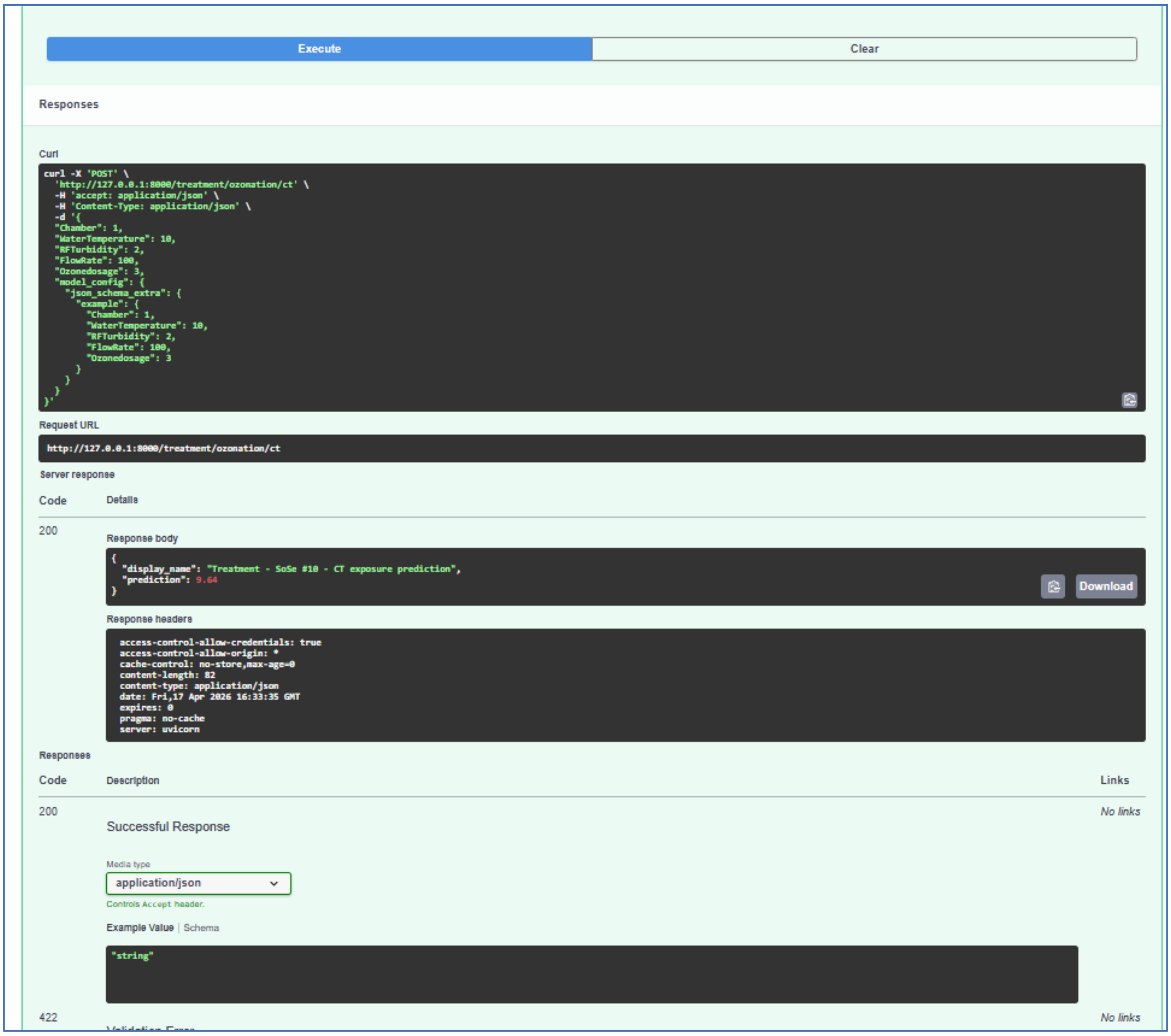
The result demonstrates the operational implementation of the coagulation soft sensor (SoSe #8) as an API-accessible service within the platform.

3.8.4.3 Ozonation stage (Treatment module)

The ozonation stage is implemented as a sequence of interconnected models within the Treatment module. First, the ozone exposure model calculates the CT value, representing the disinfection capacity of the ozonation process. This output is then passed to:

- the bromate formation model, which predicts disinfection by-product formation,
- the coliform removal model, which predicts microbiological removal efficiency.

This sequential execution demonstrates how multiple prediction services operate in combination within the same module, capturing the interaction between disinfection performance, by-product formation, and microbial safety within a unified backend workflow.



```
curl -X 'POST' \
  'http://127.0.0.1:8000/treatment/ozonation/ct' \
  -H 'accept: application/json' \
  -H 'Content-Type: application/json' \
  -d '{
  "Chamber": 1,
  "WaterTemperature": 10,
  "RFTurbidity": 2,
  "FlowRate": 100,
  "Ozonedosage": 3,
  "model_config": {
    "json_schema_extra": {
      "example": {
        "Chamber": 1,
        "WaterTemperature": 10,
        "RFTurbidity": 2,
        "FlowRate": 100,
        "Ozonedosage": 3
      }
    }
  }
}'
```

```
http://127.0.0.1:8000/treatment/ozonation/ct
```

```
{
  "display_name": "Treatment - SoSe #10 - CT exposure prediction",
  "prediction": 9.64
}
```

```
 access-control-allow-credentials: true
 access-control-allow-origin: *
 cache-control: no-store,max-age=0
 content-length: 82
 content-type: application/json
 date: Fri,17 Apr 2026 16:33:35 GMT
 expires: 0
 pragma: no-cache
 server: uvicorn
```

Successful Response

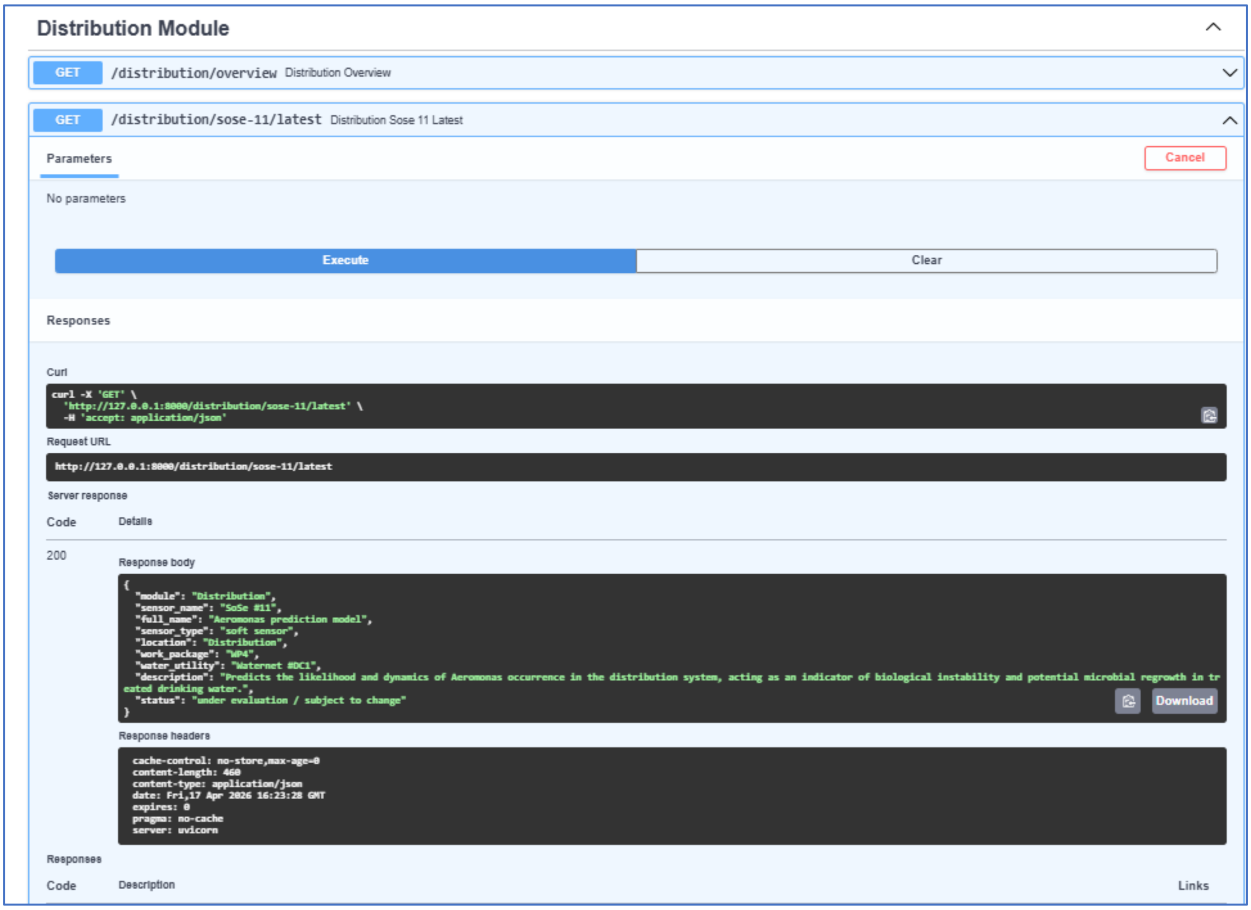
```
"string"
```

Figure 31: Example execution of the ozonation CT prediction endpoint within the Treatment module

The response indicates temporary unavailability of the model due to a missing dependency, demonstrating error handling and transparency of the API in communicating model status.

3.8.4.4 Distribution stage (Distribution module)

At the distribution level, the platform incorporates a microbiological prediction model (e.g., Aeromonas) within the **Distribution module**, as an indicator of biological water quality and potential instability in the network. This module is designed to be extendable and may evolve further based on end-user requirements.



```
curl -X 'GET' \
  'http://127.0.0.1:8000/distribution/sose-11/latest' \
  -H 'accept: application/json'
```

```
http://127.0.0.1:8000/distribution/sose-11/latest
```

```
{
  "module": "Distribution",
  "sensor_name": "SoSe #11",
  "full_name": "Aeromonas prediction model",
  "sensor_type": "soft sensor",
  "location": "Distribution",
  "work_package": "WP4",
  "water_utility": "Waternet #DC1",
  "description": "Predicts the likelihood and dynamics of Aeromonas occurrence in the distribution system, acting as an indicator of biological instability and potential microbial regrowth in tr
eated drinking water.",
  "status": "under evaluation / subject to change"
}
```

```
 cache-control: no-store,max-age=0
 content-length: 460
 content-type: application/json
 date: Fri,17 Apr 2026 16:23:28 GMT
 expires: 0
 pragma: no-cache
 server: uvicorn
```

Figure 32: Example execution of the Distribution module endpoint, showing the retrieval of the latest output from a microbiological soft sensor (SoSe #11)

The Aeromonas-related model is presented as an illustrative example, as its final implementation is still under evaluation. The figure demonstrates that the Distribution module is implemented and capable of hosting prediction services, providing the necessary structure for future integration of distribution-level sensors.

3.8.4.5 Final platform output (Risk module)

The outputs of all modules are transmitted to the **Risk module**, where they are aggregated and evaluated. The Risk module assesses threshold exceedance, assigns risk levels, and generates operational alerts and mitigation recommendations.

By combining the outputs of all models, the platform generates:

- water quality indicators,
- process performance indicators,
- disinfection performance metrics,
- by-product risk information,
- operational recommendations,
- alerts and decision-support outputs.

The final backend output therefore represents not only a set of predictions, but an integrated interpretation of system performance through risk categorization and suggested actions.

To demonstrate the implementation of the risk assessment functionality within the platform, presents the FastAPI Swagger interface of the **Risk module**. Figure 33 illustrates the available API endpoints that enable risk evaluation at different levels, including individual soft sensors, treatment-stage assessment, and overall source-to-tap risk analysis.

These endpoints receive inputs from the prediction modules (Source, Treatment, and Distribution) and provide structured outputs in the form of risk indicators, alarm flags, and suggested operational actions. This confirms that the Risk module is not only conceptually defined, but also implemented as an operational component within the backend, capable of aggregating model outputs and supporting decision-making.

Overall, the figure highlights the modular and interoperable design of the platform, where predictive models and risk evaluation are integrated through API-based communication.

Risk Module

GET `/risk/overview` Risk module overview
POST `/risk/source/sose-9/current` Source SoSe #9 risk assessment
POST `/risk/coagulation` Coagulation risk assessment
POST `/risk/treatment/rl-5/current` RL #5 dosage risk assessment
POST `/risk/ct` CT risk assessment
POST `/risk/bromate` Bromate risk assessment
POST `/risk/coliform` Coliform-removal risk assessment
POST `/risk/ozonation/current` Combined ozonation risk
POST `/risk/distribution/sose-11/current` Distribution SoSe #11 risk assessment
POST `/risk/treatment/current` Current treatment risk
GET `/risk/system/current` Current system-wide risk placeholder
POST `/risk/system/current` Current system-wide risk

Figure 33: Click on the risk module.

To illustrate the implementation of the risk assessment logic within the platform, presents an example of the response generated by the Risk module. Specifically, the SoSe 9 is chosen in the Figure 34, which demonstrates how prediction outputs from the Source module are translated into structured risk information through the backend.

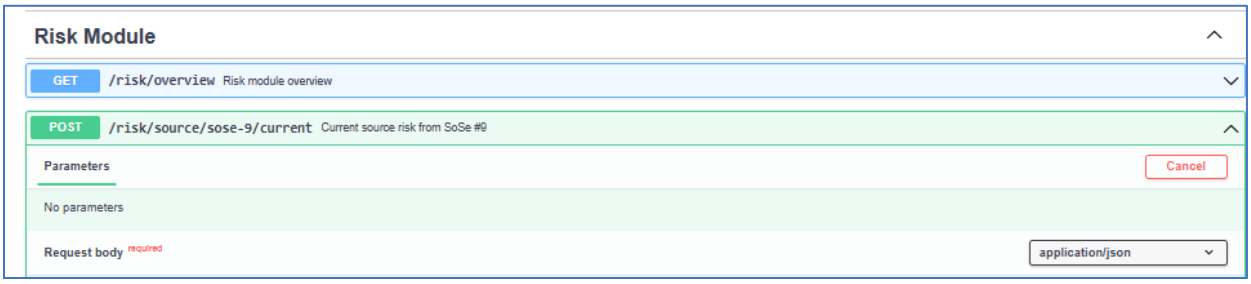
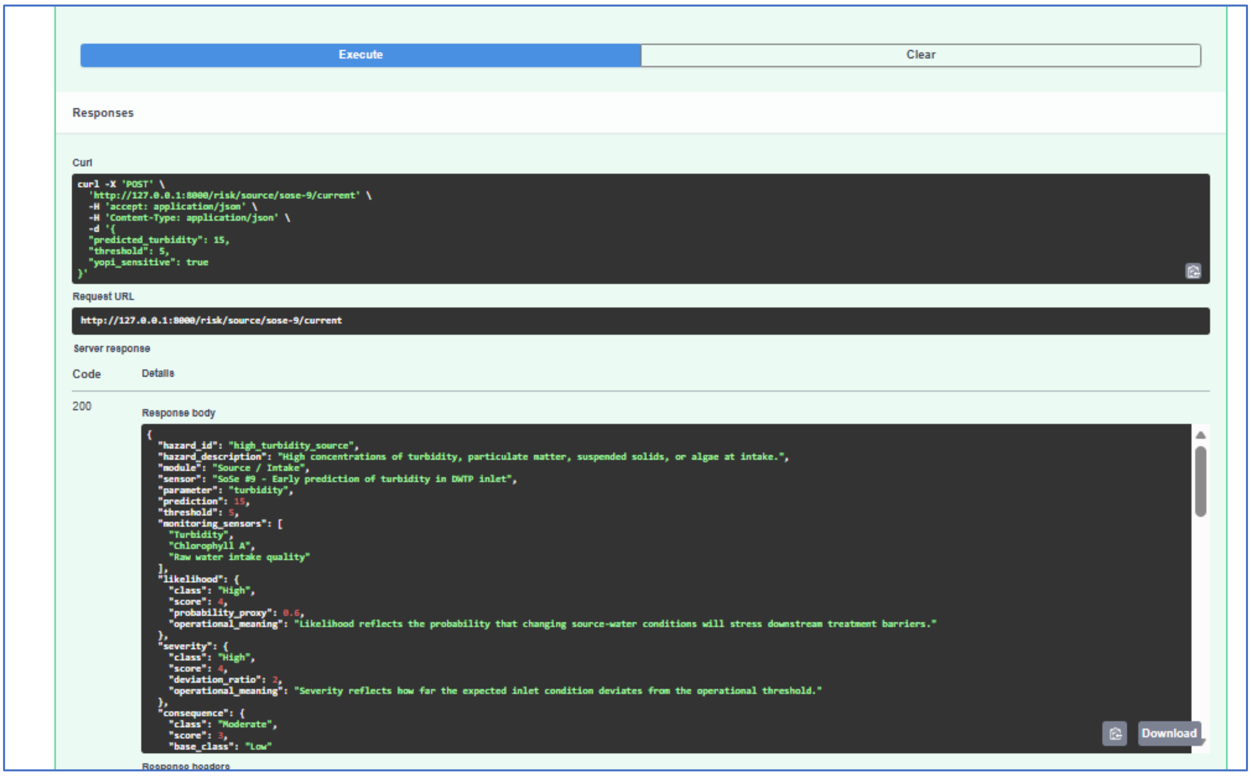


Figure 34: Choose for example the SoSe #9 which predicts the turbidity at the inlet.

The subsequent figure, shows the response body after clicking on the Execute button (Figure 35). The presented example is based on simulated input data and preliminary threshold values. These thresholds, along with the associated mitigation strategies, are currently under refinement in collaboration with end-users. Nevertheless, the figure highlights that the full risk assessment architecture has been implemented, enabling the transformation of identified hazards, described in Section 3.2 and Table 4, into executable code within the platform.

In particular, the response structure reflects the methodological framework introduced earlier, as it includes explicit representations of likelihood, severity, and consequence, along with probability proxies and operational interpretations. This demonstrates the integration of the conceptual risk framework into a functional decision-support system.



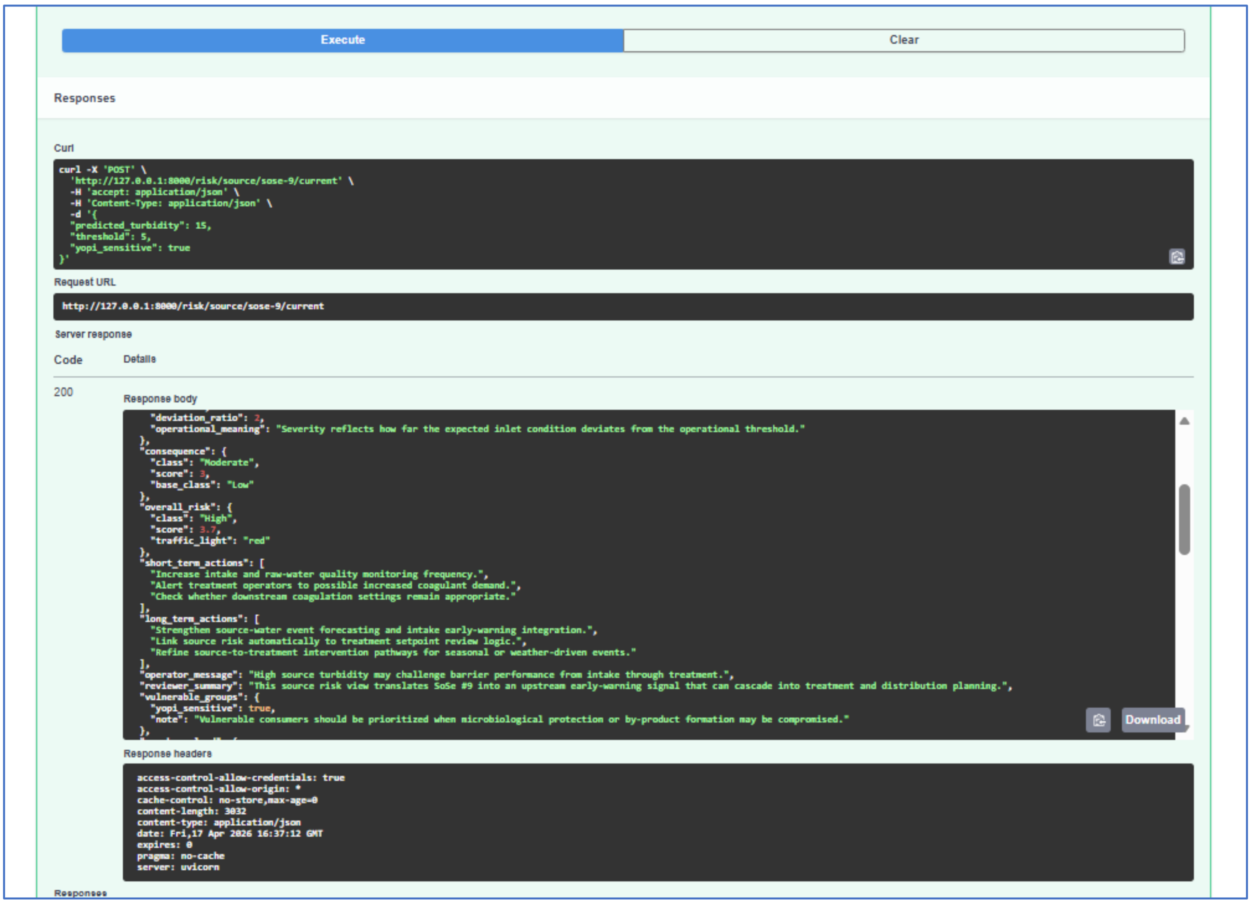
```
curl -X 'POST' \
  'http://127.0.0.1:8000/risk/source/sose-9/current' \
  -H 'accept: application/json' \
  -H 'Content-Type: application/json' \
  -d '{
  "predicted_turbidity": 15,
  "threshold": 5,
  "yopi_sensitive": true
}'
```

```
http://127.0.0.1:8000/risk/source/sose-9/current
```

```
{
  "hazard_id": "high_turbidity_source",
  "hazard_description": "High concentrations of turbidity, particulate matter, suspended solids, or algae at intake.",
  "module": "Source / Intake",
  "sensor": "SoSe #9 - Early prediction of turbidity in DWTP inlet",
  "parameter": "turbidity",
  "prediction": 15,
  "threshold": 5,
  "monitoring_sensors": [
    "Turbidity",
    "Chlorophyll A",
    "Raw water intake quality"
  ],
  "likelihood": {
    "class": "High",
    "score": 4,
    "probability_proxy": 0.6,
    "operational_meaning": "Likelihood reflects the probability that changing source-water conditions will stress downstream treatment barriers."
  },
  "severity": {
    "class": "High",
    "score": 4,
    "deviation_ratio": 2,
    "operational_meaning": "Severity reflects how far the expected inlet condition deviates from the operational threshold."
  },
  "consequence": {
    "class": "Moderate",
    "score": 3,
    "base_class": "Low"
```

Figure 35: Example response from the Risk module, illustrating the transformation of model outputs into structured risk information

The example is based on simulated data, while threshold values and mitigation strategies are currently under refinement. The response includes likelihood, severity, and consequence components, as defined in Section 3.2, demonstrating the implementation of the conceptual risk framework into executable backend logic (Part 1).



```
curl -X 'POST' \
  'http://127.0.0.1:8000/risk/source/sose-9/current' \
  -H 'accept: application/json' \
  -H 'Content-Type: application/json' \
  -d '{
  "predicted_turbidity": 15,
  "threshold": 5,
  "yopi_sensitive": true
}'
```

```
http://127.0.0.1:8000/risk/source/sose-9/current
```

```
    "deviation_ratio": 2,
    "operational_meaning": "Severity reflects how far the expected inlet condition deviates from the operational threshold."
  },
  "consequence": {
    "class": "Moderate",
    "score": 3,
    "base_class": "Low"
  },
  "overall_risk": {
    "class": "High",
    "score": 3.7,
    "traffic_light": "red"
  },
  "short_term_actions": [
    "Increase intake and raw-water quality monitoring frequency.",
    "Alert treatment operators to possible increased coagulant demand.",
    "Check whether downstream coagulation settings remain appropriate."
  ],
  "long_term_actions": [
    "Strengthen source-water event forecasting and intake early-warning integration.",
    "Link source risk automatically to treatment setpoint review logic.",
    "Refine source-to-treatment intervention pathways for seasonal or weather-driven events."
  ],
  "operator_message": "High source turbidity may challenge barrier performance from intake through treatment.",
  "reviewer_summary": "This source risk view translates SoSe #9 into an upstream early-warning signal that can cascade into treatment and distribution planning.",
  "vulnerable_groups": {
    "yopi_sensitive": true,
    "note": "Vulnerable consumers should be prioritized when microbiological protection or by-product formation may be compromised."
  },
```

```
access-control-allow-credentials: true
access-control-allow-origin: *
cache-control: no-store,max-age=0
content-length: 3032
content-type: application/json
date: Fri,17 Apr 2026 16:37:12 GMT
expires: 0
pragma: no-cache
server: uvicorn
```

Figure 36: Continuation of the response body. Part 2 of the Figure 35

As already mentioned, the current output is based on **synthetic (random) data**, but it follows the full methodological framework defined for the platform. A predicted turbidity value of 15 NTU, exceeding the operational threshold (under discussion, here it is chosen 5NTU for demonstration purposes), results in a **high likelihood (score = 4)** due to the increased probability of stressing downstream treatment processes, and a **high severity (score = 4)** reflecting the magnitude of deviation from acceptable conditions. These components are combined with consequence considerations to derive an **overall high risk classification (traffic light: red)**. The output further includes short- and long-term mitigation actions, operator-oriented messages, and a decision-tree structure, demonstrating how the platform operationalizes the Water Safety Plan (WSP) approach by linking sensor predictions to hazard identification, risk quantification, and intervention pathways from source to tap, with explicit attention to vulnerable consumer groups.

```
"short_term_actions": [
    "Increase intake and raw-water quality monitoring frequency.",
    "Alert treatment operators to possible increased coagulant demand.",
    "Check whether downstream coagulation settings remain appropriate."
  ],
  "long_term_actions": [
    "Strengthen source-water event forecasting and intake early-warning integration.",
    "Link source risk automatically to treatment setpoint review logic.",
    "Refine source-to-treatment intervention pathways for seasonal or weather-driven events."
  ],
  "operator_message": "High source turbidity may challenge barrier performance from intake through treatment.",
  "reviewer_summary": "This source risk view translates SoSe #9 into an upstream early-warning signal that can cascade into treatment and distribution planning.",
  "vulnerable_groups": {
    "yopi_sensitive": true,
     "note": "Vulnerable consumers should be prioritized when microbiological protection or by-product formation may be compromised."
  },
```

Figure 37: Short and long term actions (output from the risk assessment module

```
  "decision_tree": {
    "root_question": "Is the current Source / Intake condition acceptable?",
    "if_no": {
      "hazard": "high_turbidity_source",
      "risk_class": "High",
      "short_term_actions": [
        "Increase intake and raw-water quality monitoring frequency.",
        "Alert treatment operators to possible increased coagulant demand.",
        "Check whether downstream coagulation settings remain appropriate."
      ],
      "long_term_actions": [
        "Strengthen source-water event forecasting and intake early-warning integration.",
        "Link source risk automatically to treatment setpoint review logic.",
        "Refine source-to-treatment intervention pathways for seasonal or weather-driven events."
      ]
    },
    "if_yes": {
      "message": "Continue routine monitoring and trend analysis."
    }
  }
```

Figure 38: Decision tree - output of the risk assessment module

3.8.5 Demonstration setup and dummy data implementation

During development, direct access to WTNT’s PIMS was not available due to security and confidentiality restrictions. For this reason, a demonstration setup was implemented using dummy data. The dummy data source simulates real operational inputs and allows the backend to execute the complete workflow from data acquisition to prediction, risk assessment, and output generation.

A mock PIMS layer was created for this purpose. For example, in the coagulation case, the mock input includes raw water turbidity, ferric chloride concentration, temperature, pH, treatment lane, and time-related variables. These values are exposed through a dedicated API route, enabling the backend and the user interface to operate as if they were connected to a live system .

This design ensures that:

- the format of the data remains compatible with the future real PIMS connection,
- the prediction logic remains unchanged,
- testing and validation of the entire platform can proceed independently of external systems.

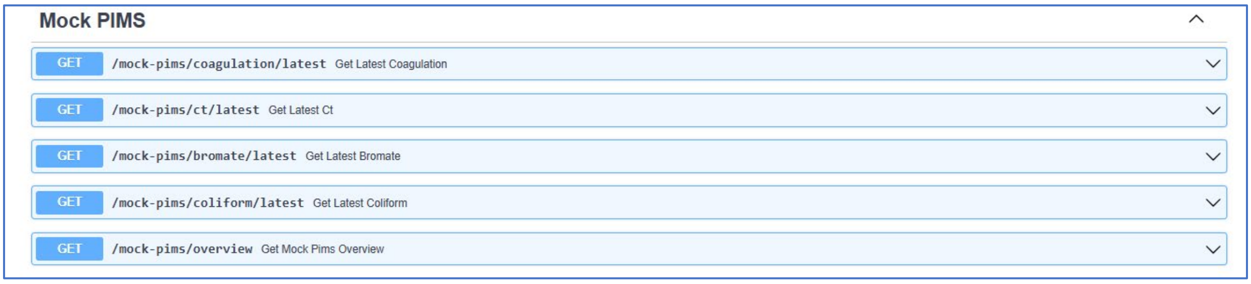
3.8.6 Live prediction and user interface support

To demonstrate real-time-like operation, a live prediction workflow has been implemented for selected models, currently exemplified through the coagulation prediction case. In this setup, the platform periodically retrieves the latest values from the mock PIMS endpoint, transforms them into the required model input structure, executes the prediction service, and returns the results together with a timestamp.

The user interface has been extended to support both:

- **manual mode**, where users enter values directly,
- **live mode**, where values are retrieved automatically from the mock backend source.

In live mode, the interface displays updated input values, prediction results, and a rolling time-series chart, allowing stakeholders to observe how predictions evolve over time and how the backend behaves in a realistic operational scenario. Figure 39 presents the graphical user interface of the platform, where both manual and live prediction modes are available. The interface allows users to visualize input data, model predictions, and system behavior over time, supporting operational decision-making in a user-friendly environment.



Mock PIMS

GET `/mock-pims/coagulation/latest` Get Latest Coagulation

GET `/mock-pims/ct/latest` Get Latest Ct

GET `/mock-pims/bromate/latest` Get Latest Bromate

GET `/mock-pims/coliform/latest` Get Latest Coliform

GET `/mock-pims/overview` Get Mock Pims Overview

Figure 39: It presents the graphical user interface of the platform, where both manual and live prediction modes are available

The interface allows users to visualize input data, model predictions, and system behavior over time, supporting operational decision-making in a user-friendly environment.

3.8.7 Transition to real deployment at WTNT

The current demonstration mode is intentionally designed so that only the data source needs to change in the final operational deployment. In the final WTNT configuration:

- the dummy data source will be removed or disabled,
- the existing WTNT connector will retrieve real-time values from PIMS,
- these values will be passed to the same API endpoints already implemented in the backend,
- the model execution logic and the user interface logic will remain unchanged.

This design ensures a seamless transition from demonstration to operational deployment. Because the platform is API-based and containerized, the backend can be deployed directly into WTNT's infrastructure as a Docker image, without interfering with the existing SCADA/DCS environment .

3.8.8 FIWARE alignment, NESSIE integration, and AWS deployment

The platform has been developed in a FIWARE-aligned manner, meaning that it supports standardized API communication, modular architecture, and interoperability with broader digital ecosystems. Although the current implementation uses direct API calls, the architecture is compatible with FIWARE principles and can be extended toward NGSI-LD-based communication.

A key requirement of the project is that the platform should not only function within the WTNT environment, but also be represented at project level within the NESSIE environment.

3.8.8.1 WTNT deployment:

The Docker image is deployed within the WTNT environment and connected to real operational data through PIMS.

3.8.8.2 NESSIE/project-level deployment:

The same Docker image is also uploaded to a cloud environment, using the Amazon Web Services (AWS) account available for the project. From there, it can be accessed through API requests and connected with NESSIE. In the final NESSIE environment, two case studies will be represented:

- the EYDAP case study, including the functionalities developed by NTUA,
- the WTNT case study, connected to the TU Delft Docker backend and including the relevant soft sensors and applicable hard sensors for the WTNT use case.

Due to confidentiality restrictions, the NESSIE/cloud deployment uses dummy data for the WTNT case where required, while real operational data remain protected within the WTNT environment. In this way, NESSIE becomes the final integrated project-level environment in which both case studies can be represented within a unified platform framework.

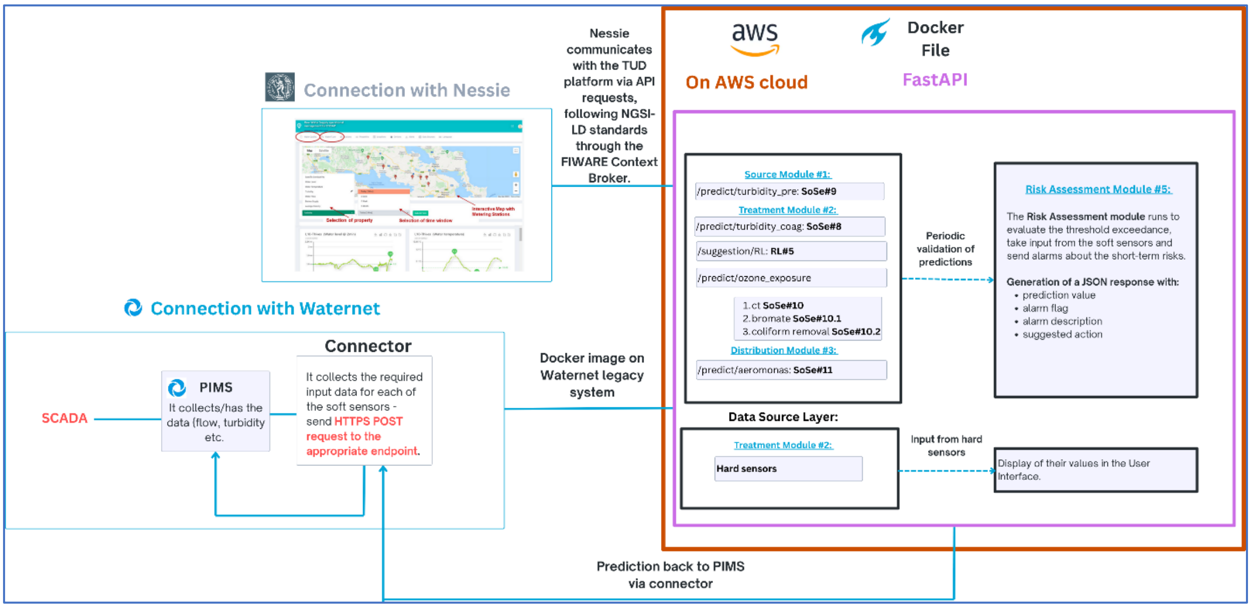


Figure 40: Deployment and integration architecture of the ToDrinQ platform, showing the connection with WTNT through PIMS and a dedicated connector, as well as the cloud-based deployment on AWS enabling interoperability with the NESSIE platform

3.8.9 Running the platform locally

To support reproducibility and future use of the platform, this section provides brief instructions for downloading and running the ToDrinQ platform locally. The platform is implemented as a FastAPI backend and can be launched from a local development environment once the project files and dependencies are available.

The steps below describe the basic procedure for starting the platform and opening the Swagger user interface.

Step 1. Download or clone the project

First, download the project folder from the repository or clone it from GitHub to a local directory on the computer.

For example, the project may be stored locally in a folder such as:

C:\Users\YourName\todrinq-platform\soft_sensors_api

The exact path may differ depending on the user’s computer and chosen folder location.

Step 2. Open PowerShell

Open **Windows PowerShell** and navigate to the folder where the FastAPI application is stored.

For example:

```
Set-Location "C:\Users\YourName\todrinq-platform\soft_sensors_api"
```

This command moves the terminal to the correct project directory.

Step 3. Activate the virtual environment

Activate the Python virtual environment in which the required packages are installed:

```
& ".\.venv\Scripts\Activate.ps1"
```

If the environment is activated successfully, the terminal will usually show the environment name at the beginning of the command line, for example:

```
(.venv)
```

Step 4. Run the FastAPI application

Start the platform with the following command:

```
uvicorn app.main:app --reload
```

This launches the FastAPI backend locally. The --reload option allows the server to restart automatically when code changes are made, which is useful during development and testing.

Step 5. Open the Swagger interface

Once the application is running, open a web browser and go to:

http://127.0.0.1:8000/docs

This opens the **Swagger UI**, where all implemented API endpoints can be viewed and tested.

Step 6. Explore the platform modules

Within the Swagger interface, the user can access the implemented modules of the platform, including:

- **Source module**
- **Treatment module**
- **Distribution module**
- **Risk module**

Each module contains its corresponding API endpoints, which can be expanded and executed directly through the interface.

Step 7. Stop the platform

To stop the local server, return to the PowerShell window and press:

Ctrl + C

This terminates the running FastAPI application.

Important note

The exact local folder path depends on where the user stores the repository on their own computer. Therefore, the example path shown above should be replaced with the actual path of the downloaded project.

In addition, some endpoints may depend on local model files, mock data, or specific package versions. For this reason, the platform should be run together with the complete project structure and required dependencies as provided in the repository.

4. Conclusions

4.1 Identification of user needs

It may be concluded that it is the task of the operator to obtain a maximum performance of the installed infrastructure, using the available information. Advanced systems in Europe rely a lot on stability and robustness. Automated systems are used for e.g. dosing of chemicals based on real-time water quality parameters. Although, they address real-time water quality changes effectively they are usually limited by scope and interdependencies, leading to sub-optimal operation. Further integration of dynamic, multi-parameter model control could improve operations.

In addition, drinking water supply (like at WTNT and EYDAP) can face (temporary) changes in source and thus source water quality. Operators should therefore have tools to support their operational decisions in terms of intake, conveyance, treatment and distribution. In addition, online visualisation of incoming water quality and predictions in the drinking water supply system as a whole and individual processes in particular are relevant for the operators to act adequately (in a risk-based manner).

Dynamic operation of e.g. the DWT processes can lead to optimization in terms of costs and water quality and minimization of the associated risks. Examples are :

- Extending run time of activated carbon to make maximal use of the installed adsorption capacity, while having sufficient capacity to adequately remove emerging compounds such as PFAS.
- Optimising the dosage of coagulants and disinfectants (chlorine and ozone) to be cost-effective, diminishing the DBP formation, with associated health risks, while avoiding clogging and maintaining sufficient disinfection capacity.
- Controlling re-growth in distribution networks with associated health risks and discoloration events, by maintaining low AOC and biofilm growth potential levels in drinking water in non-chlorinated water.

However, for dynamic operation of e.g. DWT plants adequate information is needed, such as influent quality from the source and conveyance system, being able to anticipate on changes in nutrient loads, turbidity levels, concentrations of pathogenic micro-organisms and emerging OMPs, such as PFAS. That information is scarcely (real-time) available at the production location, to be able to optimize operation of a DWT plant dynamically on multiple parameters.

Data are collected, but are difficult to access, also because they are stored at various platforms (LIMS, SCADA, PRODIS, RTPM, off-line log books). These data are, furthermore, hardly processed to information for operators (e.g. by combining data, model and predictions), and historical trends of operation are not always present. The available data are, if possible, mainly used by process technologists, who translate information into static operational rules. For transforming data to information, it is of importance that, on the one hand, the process conditions are known when (water quality) data are collected, preferably real-time. Therefore, it is needed to know beforehand which information is needed for dynamic process operation to be able to decide on the necessary data collection and processing.

4.2 Contribution to monitoring and operational support

WP7 of the ToDrinQ, focuses on the development of an integrated modular platform for Drinking Water Systems' monitoring and operational support. Based on the identified risks for the drinking water quality in the drinking water supply system at WTNT and EYDAP a conceptual design is made for the various modules: Source, Conveyance, Treatment, Distribution. The modules are based on data models that have an input of online water quality and quantity data and operational parameters. The data and parameters come from conventional sensors, present in the actual water supply system, extended with data from hard sensors (e.g. pathogenic micro-organisms, nutrients and heavy metals) developed within ToDrinQ (WP3) and soft-sensors (e.g. bromate formation, AOC formation, algae blooms, WQI), developed within ToDrinQ (WP4). The output of the modules is again data on water quality and quantity, obtained by hard and soft sensors. The sub-modules (e.g. coagulation, ozonation, activated carbon filtration within the treatment module) and entities within the module are chosen based on the priorities of the end-users WTNT and EYDAP. Within the module decision support is given by the Reinforced Learning models developed within ToDrinQ (WP4), e.g. required coagulant dose, required ozone dose, to avoid risk of degradation of the water quality within a module. Risks have also been identified that are overarching the individual modules and require a holistic operational approach, leading to some alternative pathways to mitigate the risks.

4.3 Innovation and pathways to impact

Before the ToDrinQ project, open access tools to support operators on drinking water supply to maintain a high drinking water quality did not exist. Therefore, the needs of end-users have been identified and the conceptual design of the platform and the operability standards were set. We started from scratch, but during the duration of the project we expect to reach TRL 5. We do not envision a patent and want to develop an open access tool that can be used by a wide range of (end-)users. The platform will make use of FIWARE and NESSIE and in such way the open use of the platform is guaranteed. We will first test the developed platform with the end-users within the ToDrinQ project WTNT and EYDAP. When at reasonable TRL the platform will be disseminated through the Zeropollution4Water and the ICT4Water clusters, making more end-users interested with possibilities for upscaling.

The revised EU Drinking Water Directive promotes a risk assessment and risk management approach for securing drinking water supply in the context of climate change and increased pollution. However, this approach is challenged by insufficient information that is available to operators, especially in real time, on compounds and organisms of emerging concern, such as pesticides, pharmaceuticals, disinfection by-products, heavy metals and pathogenic microorganisms. We argue that if drinking water treatment could leverage novel technologies and design philosophies, and more agile operational actions could be supported, drinking water supply systems could become more adaptable and robust without expensive infrastructural investments. In this context, ToDrinQ develops and tests a compendium of modular, complementary, innovative solutions (the 'ToDrinQ Toolkit') that provide new information and better support tools to operators and designers to adapt to (short- and long-term) changes in water quality, while obtaining high drinking water quality at the tap. ToDrinQ develops novel real time sensing and water quality monitoring technologies, innovative treatment systems (especially suitable for small-scale/modular, adaptable treatment plants) and interoperable decision tools that support resilient, evidence-based treatment plant design and improved overall water system operational awareness and response.

ToDrinQ

Funded by
the European Union