Ref. Ares(2024)1498407 - 27/02/2024



Monitoring system deployment plan

Deliverable 3.2

WP 3 DISTRIBUTION – Advanced Monitoring and Prediction

Author: Christoph Wagner
& Pavlos Pavlou

Date: 26.02.2024

Funded by
the European Union

GRANT AGREEMENT NUMBER	101081728		
FULL TITLE / ACRONYM	INNOVATIVE TOOLS TO CONTROL ORGANIC MATTER AND DISINFECTION BYPRODUCTS IN DRINKING WATER / intoDBP		
START DATE	01.12.2022	**DURATION**	48 months
END DATE	30.11.2026		
PROJECT URL	**https://intodbp.eu/**		
DELIVERABLE	**D3.2**		
WORK PACKAGE	**WP3**		
CONTRACTUAL DATE OF DELIVERY	**29.02.2024**		
ACTUAL DATE OF DELIVERY	**29.02.2024**		
NATURE	**REPORT**	**DISSEMINATION LEVEL**	**PUBLIC**
LEAD BENEFICIARY	**SCAN**		
RESPONSIBLE AUTHOR	Christoph Wagner		
CONTRIBUTIONS FROM	SCAN, WBL, UCY		
ABSTRACT	This deliverable consists of the design of the monitoring plan for the case study of the Limassol Water Distribution Network according to the monitoring technologies described and provided by WP2. A methodology for optimal sensor location along with the required tools to implement a holistic monitoring system in the CS are identified and communicated to the CS owners.		

Disclaimer

Any dissemination of results reflects only the author's view, and the European Commission is not responsible for any use that may be made of the information it contains.

Copyright message

© intoDBP Consortium, 2024
This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation, or both. Reproduction is authorised provided the source is acknowledged.

TABLE OF CONTENTS

LIST OF TABLES 4

LIST OF FIGURES 4

EXECUTIVE SUMMARY 5

1.0 LIMASSOL CASE STUDY DESCRIPTION 6

1.1 General information 6
1.2 Case study 6

2.0 SENSOR INFORMATION 9

2.1 Types of installed sensors and configuration 9
2.2 Technical specifications 10
2.3 Sensors to be installed in the future 11

3.0 OPTIMAL SENSOR PLACEMENT METHODOLOGY 11

3.1 Description of the used methodology 11
3.2 Application to the case study location 13

4.0 SETTING UP THE PILOT 16

4.1 Sensor installation locations 16
4.2 Construction, telecommunication, and other works 17

5.0 INFORMATION COMMUNICATION TECHNOLOGY - ARCHITECTURE PLAN 19

5.1 Data acquisition system plan and ICT diagram 19
5.2 Data visualization and retrieval – User interface 20

6.0 INTEGRATED MONITORING PLAN 25

6.1 Connection with tasks T3.3 and T3.4 25

LIST OF TABLES

Table 1: Summary of s::can sensor specifications 10
Table 2: Explanation of the different folders on the intodbp.s-can.at data server 19

LIST OF FIGURES

Figure 1: intoDBP case study area of the Water Distribution Network of Limassol, Cyprus. 7
Figure 2 : a) Zone 2 tank's booster chlorinator using sodium hypochlorite dosing, b) Zone 2 tank structure c) Grafana user – interface for SCADA data visualization. 8
Figure 3: Components the SCAN micro::station (exemplary photo) 10
Figure 4: Hydraulic model of DMA 230 using the EPANET software showing network elevation and pipe diameters (left) and average water demand (right) 14
Figure 5: Full (left) and skeletonized (right) hydraulic model of DMA 230 showing node pressures and pipe flows. 15
Figure 6: Sensor placement results. Left: Map with the locations for the sensor installation and the sampling procedure. Right: Map indicating all possible locations with high water age and consistent flow directions over time. 15
Figure 7: Map of the case study water distribution network showing the location of the installed SCAN sensor stations. 1) Outlet of the DWTP, 2) outlet of the Zone 2 tank, 3) Inlet of DMA 230, 4) Distribution point within DMA 230 17
Figure 5: Different works executed for the deployment of the monitoring system: 1) Installation of the sensors and connection with the terminal, 2) Installation of a solar panel to provide electricity for the operation of the sensor station and the transmission of data, 3) construction works and adjustment of pipes for the installation of the sensors, 4) the terminal screen showing time series data, 5) the terminal screen showing the measured parameters. 18
Figure 6: Data management plan as developed by SCAN 20
Figure 7: Set of physicochemical data (Temperature, pH, Chlorine, Conductivity) as shown through the Grafana interface 21
Figure 8: Set of turbidity and UV-Vis data as shown through the Grafana interface 21
Figure 9: Set of organic and nitrate related data as shown through the Grafana interface 22
Figure 10: Data retrieval for a selected period through the Grafana interface 22
Figure 11: Set of data retrieved between 16 -22 October filtered per parameter 23
Figure 12: Set of data retrieved between 16 -22 October filtered per location (data from the distribution location) 24
Figure 13: Flow diagram between tasks 3.2 and 3.3. The diagram explains how data retrieved from the monitoring system is used for the development of the proposed simulator of T3.3. 25

EXECUTIVE SUMMARY

This deliverable outlines the monitoring plan for the Limassol case study site. A comprehensive methodology was used for optimal sensor placement from the drinking water plant to the far end of the distribution network. The implementation of this plan involves the installation of four strategically positioned sensor stations designed to measure key parameters critical for assessing water quality and Disinfection by Product (DBP) predictions. These parameters include Total Organic Carbon (TOC), Dissolved Organic Carbon (DOC), Nitrate NO_3^-, ultraviolet-visible light (UV-Vis) spectra, turbidity, temperature, conductivity, free chlorine, and pH.

In conjunction with the sensor deployment, a robust Information Communication Technology (ICT) plan has been developed to facilitate efficient data transmission and acquisition. This plan also encompasses the creation of a user-friendly interface, ensuring accessibility and usability for stakeholders involved in the Limassol case study.

Furthermore, the methodology involves the integration of the collected sensor data into hydraulic and quality models for DBPs. These models, developed in Task 3.3, enhance the understanding of the complex interactions between hydraulic factors and water quality, contributing to a more accurate assessment of potential DBP formation. This integration provides a holistic approach to water quality management, allowing for proactive measures to mitigate the risk of DBPs and enhance overall water treatment system efficiency.

By combining strategic sensor placement, advanced ICT solutions, and model integration, this methodology not only increases the understanding of formation potential changes of DBPs but also empowers water treatment facilities with actionable insights for informed decision-making and improved water quality control.

1.0 Limassol case study description

1.1 General information

Limassol is located on the southern coast of Cyprus and is the second largest city of the island. The city experiences a Mediterranean climate characterized by hot and dry summers along with mild winters. The Water Distribution Network of Limassol (LWDN) is managed by the Water Board of Limassol (WBL) who is an intoDBP partner. The system serves 110,000 consumers (approximately 180,000 population), and the annual average water demand is estimated to be 17 million m^3. The Water Distribution Network (WDN) is separated into 9 pressure zones and includes 31 water tanks, 37 pumping stations and approximately 1,150 km of pipes with 100 km of transport pipes and 1,100 km of distribution pipes. The elevation of the supplied area ranges from 0 to 550 m above sea level. The network is supplied with water from desalination plants, groundwater, and the Limassol Drinking Water Plant (LDWP), which receives water from dams. The water that is received from the different water systems is transferred initially into water tanks managed by the WBL, re-chlorinated and then distributed for consumption.

1.2 Case study

District Metered Area (DMA) 230 is selected as the project's case study. DMA 230 falls within the pressure zone called "Zone 2". The pilot includes a transport network with 2.6 km of transport pipes that connects the LDWP with the WBL's water tank, called "Zone 2 water tank", and a distribution network with 31.2 km of pipes to distribute water from the tank to the DMA 230 (Figure 1). The selected case study covers a residential area with approximately 4663 consumers, which corresponds to a population of 12124 people, and a daily average demand of 1800 m^3. DMA 230 is located between the following 4 main streets: Leoforos Spyrou Kyprianou, Petrou Tsirou, Leoforos Archiepiskopou Makariou III, and Agias Zonis. It covers the area of postcodes 3085,3086,3087,3090, 4001 and 3083 (part of it).

The "Zone 2 water tank" mainly receives water from the LDWP, but in certain periods when there are low water reserves in the dams, water is also provided from the desalination plant. In this case, water from the LDWP and the desalination plant is mixed before entering the "Zone 2 water tank". The tank has a storage capacity of 17 000 m^3 and an average daily consumption of 12950 m^3. It serves 17 DMAs including DMA 230 with a piping system of approximately 280km length. The re-chlorination in the tank is performed by sodium hypochlorite dosing with the annual need reaching 18500 L (Figure 2).

Due to the nature and connectivity of the distribution network, initially a larger version of the case study was considered to achieve water mass balance and accurately model its hydraulic operation. This was essential to develop a calibrated model of the case study that accurately simulates the operation of the real network. The expanded version of the case study includes the two water sources (the LDWP and the desalination plant) and all the intermediate consumption areas between the source, the tank, and the selected distribution area (DMA 230) (Figure 1).

The WBL operates a Supervisory Control and Data Acquisition (SCADA) system to effectively monitor the operation of WDN. On-line field flow, pressure and level sensors are installed at the inlet and outlet of the tanks, at the entrance of each DMA and at the various pumping stations. A limited number of chlorine sensors are also installed in specific areas which are not included in this case study. In previous collaborations between the WBL and KIOS CoE, a user-interface through Grafana had already been developed to enable the visualization of existing sensor data (Figure 2). An API (Application Programming Interface) was further developed to allow the retrieval of data for various analytical purposes.

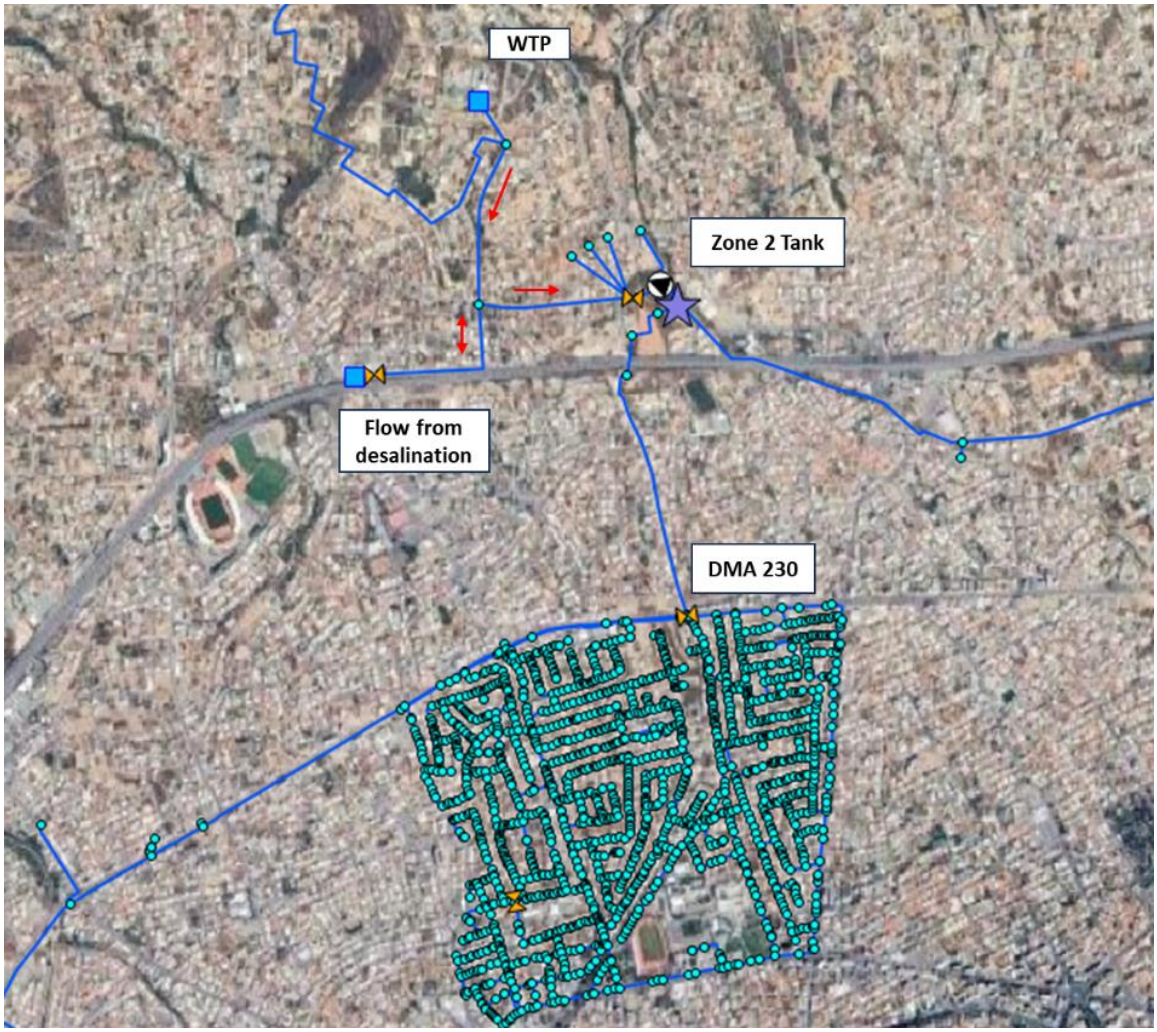


Figure 1: intoDBP case study area of the Water Distribution Network of Limassol, Cyprus.

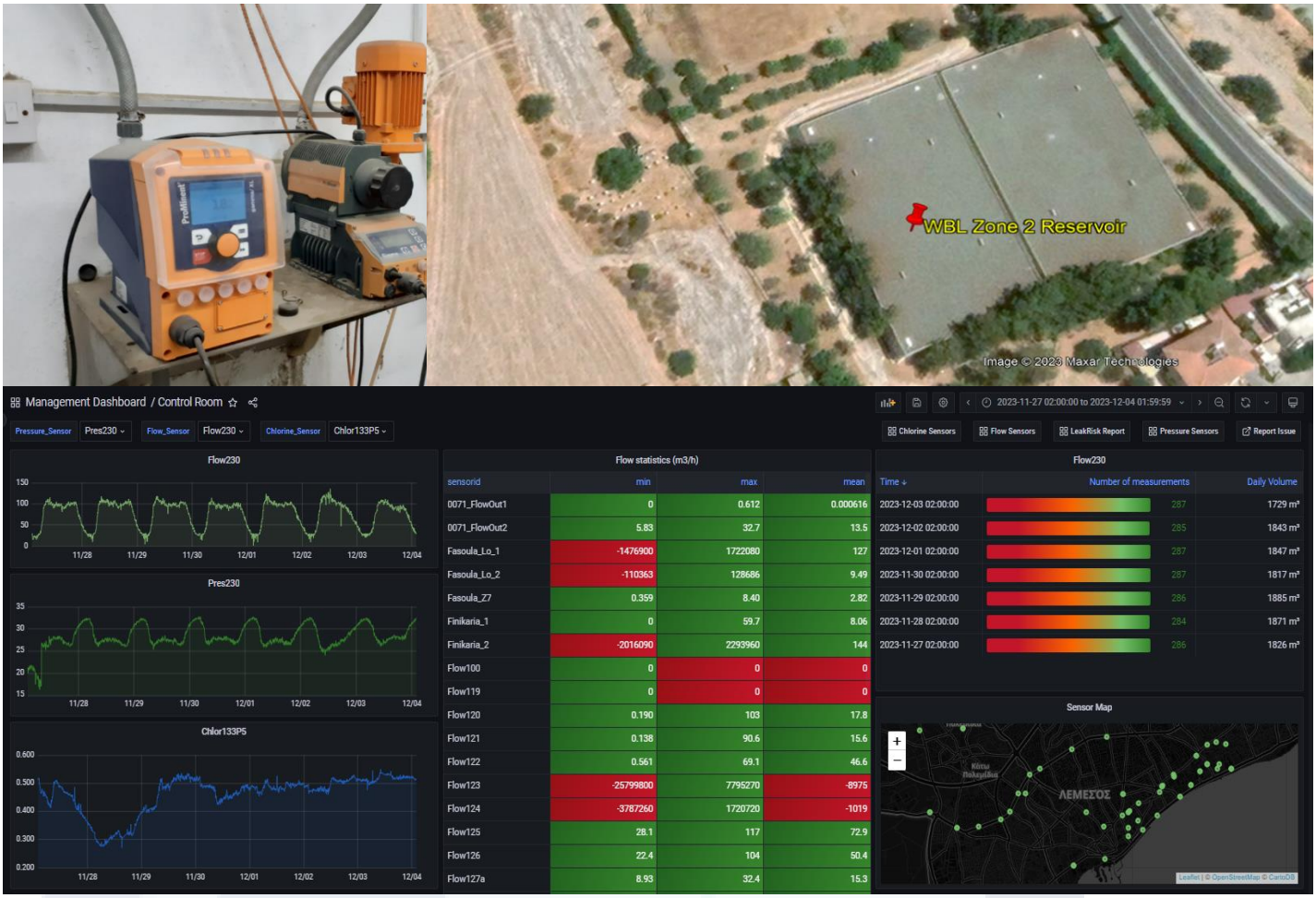


Figure 2 a) Zone 2 tank's booster chlorinator using sodium hypochlorite dosing, b) Zone 2 tank structure c) Grafana user – interface for SCADA data visualization.

2.0 Sensor information

2.1 Types of installed sensors and configuration

The measurement parameters were chosen to support a comprehensive monitoring scheme of factors influencing the formation of DBPs in the drinking water network. These factors especially include organics, free chlorine and pH. To complement the picture additional parameters such turbidity and conductivity were also added.

Each of the measurement stations installed in the Limassol case study site consist of the following sensors measuring below parameters:

- UV/VIS spectrometer (spectro::lyser)
  - Pathlength: 35 mm
  - Parameters: UV254, Turbidity, Color, DOC, TOC, NO3
  - Measurement principle: full UV/VIS spectrum using a Xenon flash lamp combined with a 256px array detector
- pH Sensor (pH::lyser)
  - Parameters: pH, Temperature
  - Measurement principle: classical glass electrode pH measurement
- Free Chlorine Sensor (chlori::lyser)
  - Parameters: Free Chlorine (FCl), Temperature
  - Measurement principle: Amperometric measurement, membrane covered, pH compensated sensor design
- Conductivity Sensor (condu::lyser)
  - Parameters: Conductivity, Temperature
  - Measurement principle: 4 electrode contacting conductivity measurement

The sensors are panel-mounted and installed in flow-through-cells. The spectrometer sits in its own flow cell also containing a brush cleaning system to keep the transmission windows clean from e.g. biofouling, which would alter the measurement. The other sensors are installed in one combined flow cell that can accommodate up to four “classical” sensors. Each panel is equipped with a metal-mesh filter to safeguard the sensors from larger particles without altering the turbidity reading. Additionally, a flow switch or a small flowmeter are included to monitor the water flow through the monitoring station, which is vital to ensure accurate FCl readings. Between the spectrometer flow cell and the sensor flow cell a flow restrictor generates a pressure drop ensuring that the spectrometer side is under the highest water pressure to reduce degassing of the aqueous sample which would lead to the formation of tiny air bubbles that potentially interfere with the optical measurement.

All sensors are fully digital, using the MODBUS RTU protocol to communicate with the control terminal. Three out of the four installed stations are equipped with touchscreen terminals called con::cube (2x 230V supply, 1x 12V DC supply). The fourth station, the most power critical one, was equipped with a newly developed low power terminal optimized for power efficiency, called con::line (distribution point in the DMA). The terminals aggregate the data and send the data over a 4G wireless communication

uplink to a centralized server (see section 5.0 for more information on data management). All settings regarding the measurement (e.g. measurement interval, cleaning cycles, data upload, etc.) are configured through the terminals directly if on premises or the can also be changed remotely through an encrypted VPN connection. Figure 3 shows an exemplary photo of such a measurement station.

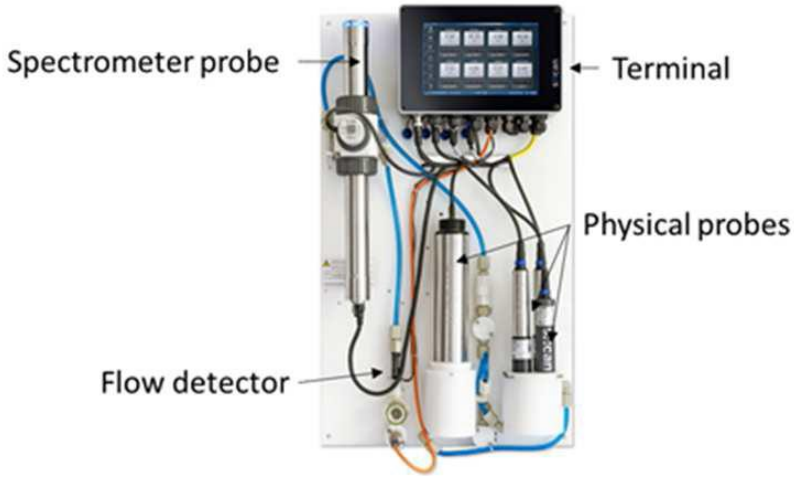


Figure 3: Components the SCAN micro::station (exemplary photo)

2.2 Technical specifications

Table 1 summarizes the technical specifications of the installed s::can sensors by parameter. Full sets of specifications can be downloaded on the s::can website[1]

Parameter	Measurement Range	Measurement Resolution	Operating Temperature	Response Time (T90)
FCl	0 – 2 mg/l	0.001 mg/l	0 – 45 °C	2 min
pH	2 – 12 pH	0.01 pH	0 – 70 °C	30 sec
Cond.	0 – 500 mS	0.01 mS	0 – 70 °C	n/a
NO_3	0 – 80 mg/l	0.01 mg/l	0 – 45 °C	n/a
DOC	0 – 17 mg/l	0.01 mg/l	0 – 45 °C	n/a
TOC	0 – 22 mg/l	0.01 mg/l	0 – 45 °C	n/a
Turb.	0 – 170 NTU	0.01 NTU	0 – 45 °C	n/a
Color app	0 – 500 Hazen	0.1 Hazen	0 – 45 °C	n/a
Color true	0 – 300 Hazen	0.1 Hazen	0 – 45 °C	n/a
UV254	0 – 70 Abs/m	0.001 Abs/m	0 – 45 °C	n/a

Table 1: Summary of s::can sensor specifications

[1] https://www.s-can.at/en/our-products/

2.3 Sensors to be installed in the future

The monitoring plan also includes the installation of 2 bioreporters and 2 fluorescence tools that will be developed in WP2 by the University of Jerusalem and the University of Catania respectively. By M24 these sensors will be integrated on the existing sensors stations.

The fluorescence sensors will be installed on two of the panels in series with the spectrometer probe by using a T-piece of piping where the water flows straight through, and the third port will be used to insert the fluorescence sensor directly into the flowing water stream. The T-piece will be installed before the flow restrictor described in Section 2.1. The readout of these sensors will be achieved through a 4-20 mA analogue signal that is digitized by the terminals and then mapped to the optical measurement range of the sensors. The data will be presented like the one from the s::can sensors and transmitted to the cloud in the same way.

The integration of the bioreporter sensors will happen through digital means, either through MODBUS TCP or MODBUS RTU. Additionally, since the bioreporter measurement is not an online measurement the terminals will trigger the start of the measurement through a digital signal. After the trigger signal the bioreporter sensor will feed water to its measurement cell and start the readout of the optical signal. The temporal evolution of this signal will be monitored by the terminal and evaluated through the “free formula” functionality of the terminal (yet to be implemented on the con::cube terminals). The delta between the baseline, right after the cell is with water, and the maximum of optical signal represents the analytical signal of interest. All data (temporal and final reading) will be transmitted to the central server as described in section 5.1.

3.0 Optimal sensor placement methodology

3.1 Description of the used methodology

The objective of the optimal sensor placement methodology, developed by KIOS, is to identify the most appropriate locations for installing the minimum number of online quality sensors aiming to detect the potential formation of DBPs and detect changes in water quality as quickly as possible. Several parameters that influence the strategy of sensor placement are identified and analysed below:

1. Pre-disinfection Water Quality

- **Key Consideration**: The composition and quality of raw water significantly impact the formation of DBPs during the disinfection process. Organic matter, bromide ions, and ammonia levels are crucial indicators.
- **Strategy**: Place sensors upstream of disinfection points to assess raw water quality and predict potential DBP formation. This enables targeted treatment adjustments to minimize DBP precursors.

2. Network Hydraulic Characteristics (Flow Patterns)

- **Key Consideration**: The dynamics of water flow within the network determine how water and any formed DBPs are distributed. Stagnant areas or zones with slow-moving water are particularly prone to higher DBP concentrations.
- **Strategy**: Install sensors in strategic locations that represent the range of flow conditions within the network, including areas prone to stagnation and high-flow conduits, to effectively monitor how flow patterns influence DBP levels.

3. Water Age - Residence Time

- **Key Consideration**: The time water spends in the distribution system before reaching consumers affects DBP levels, with longer residence times generally leading to higher concentrations.
- **Strategy**: Sensors should be placed at points across the network that represent a range of water ages, especially in parts of the system where water is known to reside longer, to assess the impact of residence time on DBP formation.

4. Water Age - Distribution

- **Key Consideration**: The distribution of water ages throughout the network can show where old and freshwater mix, which can either dilute or exacerbate DBP concentrations.
- **Strategy**: Mapping the water age distribution helps in identifying critical points for sensor placement to monitor the effects of mixing on DBP levels. This might include junctions where water from different parts of the system converges.

5. Pipe Materials

- **Key Consideration**: The materials used in the water distribution network (e.g., PVC, cast iron, copper) can interact with DBPs and influence their stability and degradation rates.
- **Strategy**: Consider the distribution of different pipe materials within the network when placing sensors, as certain areas might show different DBP behaviors. Sensors near or within areas served by pipes known to affect DBP stability can provide insights into material-related degradation or stabilization effects.

Sensor placement is mainly considered an optimization problem, assessed through either single or multi-objective approaches based on impact metrics. We aim to design a DBP sensor placement methodology with a focus on 1) maximizing network coverage and 2) enhancing the capability of DBP formation models. Here's how the methodology can be developed and implemented considering the specified metrics and objectives:

1. Metrics for Sensor Placement

- **Water Age**: This is crucial for understanding DBP formation potential, as longer water ages can lead to higher DBP concentrations due to extended reaction times.
- **Water Mixing**: Monitoring areas where old and new water mix can provide insights into how these dynamics affect DBP levels.

- **Coverage of Network Nodes**: Aiming to monitor water quality for as many consumers as possible by strategically placing sensors at nodes that represent the distribution network comprehensively.
- **Coverage of Network Pipes with Different Materials**: Since pipe materials can affect DBP stability and degradation, it's essential to cover various materials across the network.

2. Objectives for Sensor Placement

- **Increase Monitoring Coverage**: Ensure that the placement of sensors allows for a broad overview of water quality across the entire network, prioritizing areas with significant consumer populations and critical infrastructure.
- **Enhance Identification Potential of DBP Formation Models**: By selecting nodes with larger water age times and reducing path uncertainty, the methodology aims to provide clearer insights into DBP formation patterns. This is crucial for developing predictive models that can accurately forecast DBP levels under various conditions.

3. Methodology Development

a. Mapping Water Flow Paths

- **Identify all possible water flow paths** within the distribution network using hydraulic models. This step is crucial for understanding how water moves through the system, where it resides the longest, and where mixing occurs.
- **Analyze Flow Reversals**: Flow paths prone to reversals should be identified as they complicate the prediction of DBP formation due to varying water ages and mixing patterns.

b. Calculating Metrics

- Utilize hydraulic modeling tools to calculate the aforementioned metrics for each node or section of the network. This includes determining water age, identifying mixing zones, and assessing the representativeness of different network areas and pipe materials.

c. Sensor Placement Strategy

- **Prioritize Nodes**: Based on the calculated metrics, prioritize nodes for sensor placement that are critical for achieving the stated objectives. This includes nodes with long water ages, high potential for water mixing, and those that ensure broad network coverage.
- **Address Path Uncertainty**: Select nodes that are on paths with consistent flow directions over time to minimize uncertainty in DBP formation predictions.

3.2 Application to the case study location

The proposed sensor placement strategy is applied in two different cases in our case study. Firstly, it was employed to define the optimal location to place one of the sensor stations and secondly to define another extra three locations for sampling within DMA 230.

Model preparation: A hydraulic model of the distribution network of the DMA 230 was developed utilizing GIS data as described in section 6. As shown in Figure 4, the initial model includes all the area's consumers (water meter locations) thus it consists of 1998 nodes and 2068 pipe segments. Although this detailed model is considered suitable for hydraulic and quality simulations it is computationally expensive for further applications such as the sensor placement module.

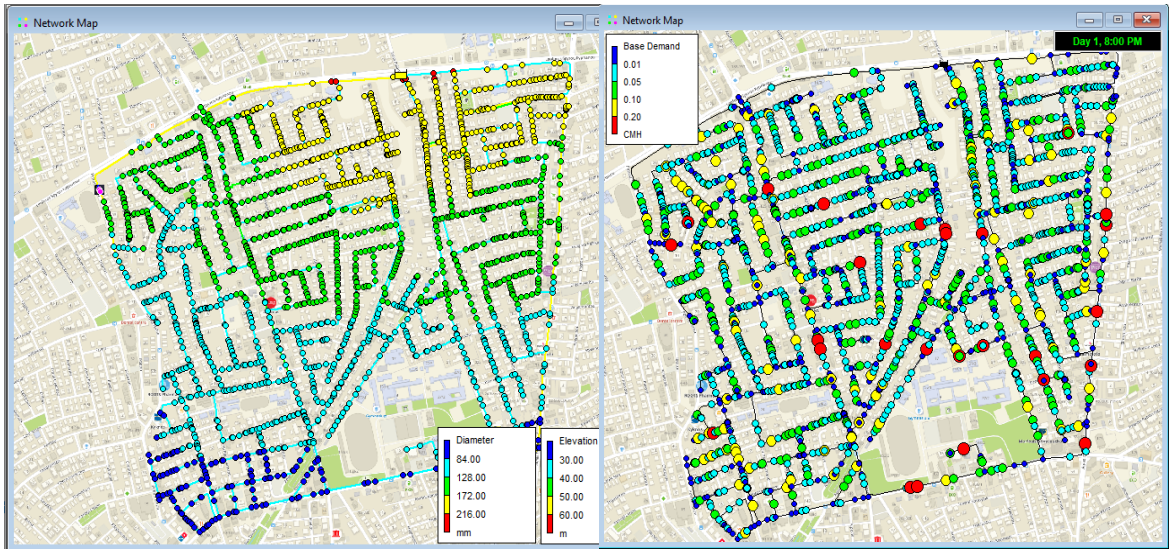


Figure 4: Hydraulic model of DMA 230 using the EPANET software showing network elevation and pipe diameters (left) and average water demand (right)

The initial hydraulic model is then processed through the skeletonization algorithm (see section 6), and a simplified model is developed (Figure 5). Both models were calibrated before and after the skeletonization process to check that the hydraulic behaviour is not affected and that the simplified model is accurate. The new simplified model has 250 nodes and 330 pipe segments which makes it suitable for further extensive analyses.

The results of the sensor placement strategy are presented in Figure 6. As previously mentioned, the proposed methodology was initially applied to detect the optimal location for 1 monitoring station. Then, the same methodology was applied to identify 3 extra locations for sampling purposes. In this case the algorithm considered the existence of 1 sensor in the network. Nodes with red text are nodes that fit the sensor placement strategy, i.e., they have paths with consistent flow directions over time and high waterage. We selected the 3 sampling locations with the aim to maximize coverage with the sensor location covering the network that does not have consistent flow paths. Regarding the sensor location, we had to consider more limitations including space availability since construction works were necessary for the installation process.



Figure 5: Full (left) and skeletonized (right) hydraulic model of DMA 230 showing node pressures and pipe flows.

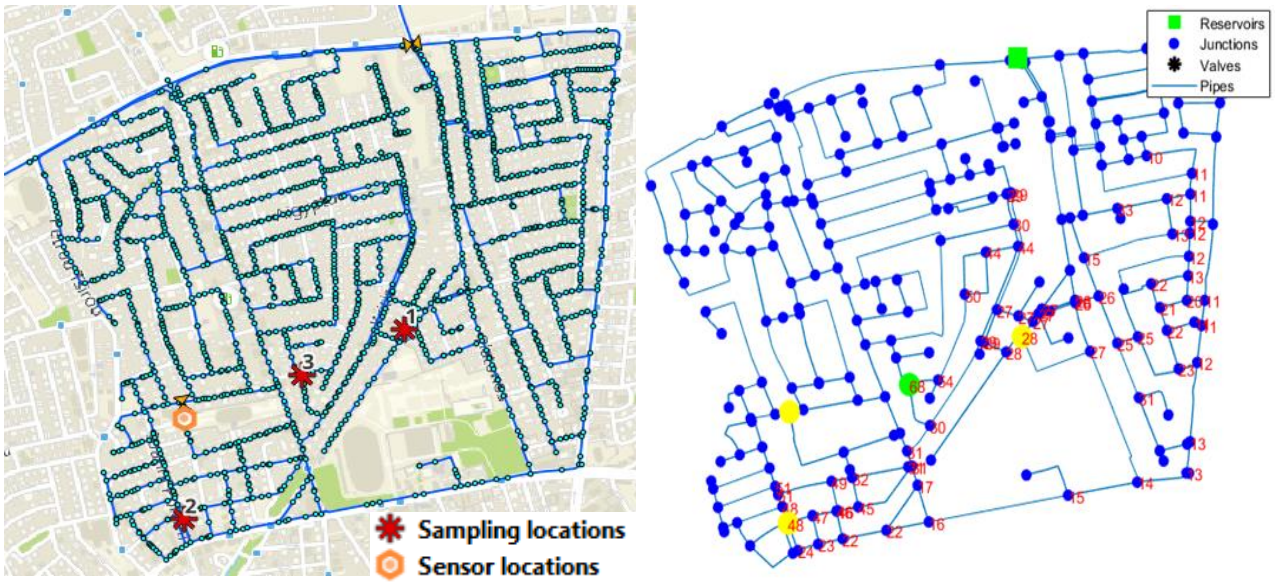


Figure 6: Sensor placement results. Left: Map with the locations for the sensor installation and the sampling procedure. Right: Map indicating all possible locations with high water age and consistent flow directions over time.

4.0 Setting up the pilot

4.1 Sensor installation locations

According to the project's plan, 4 sensor stations were installed within the selected case study area in the following locations (Figure 7):

- Outlet of the DWTP
- Outlet of the Zone 2 tank
- Inlet of the DMA 230
- Distribution point within DMA 230

The sensor stations were strategically located in these areas aiming to monitor the water quality and how it is affected between the DWTP, the tank and the distribution area.

The initial water quality status for distribution is obtained from the first sensor at the DWTP outlet, while the second sensor at the Zone 2 tank outlet offers crucial insights into the impact of re-chlorination on various quality parameters. With the tank dimensions and the booster chlorination settings known a model can be developed to simulate the re-chlorination process and calibrated using measurements from these two sensors.

Additionally, information on the fate of several DBPs and how water hydraulic properties such as water velocity, retention time in the tank, water age, and pipe material can be extracted from the data provided by the two sensor stations located at the entrance and within DMA 230. The location of the sensor station installed within DMA 230 was selected utilizing an optimal sensor placement methodology that is described in Section 3.2.



Figure 7: Map of the case study water distribution network showing the location of the installed SCAN sensor stations. 1) Outlet of the DWTP, 2) outlet of the Zone 2 tank, 3) Inlet of DMA 230, 4) Distribution point within DMA 230

4.2 Construction, telecommunication, and other works

A series of construction and other works had to be taken for the efficient installation and operation of the monitoring system. For the two sensors stations installed at the outlet of the DWTP and the Zone 2 tank, the main works were the installation of the sensors and their connection with the terminal. Since these sensors were installed in existing infrastructure and within or next to buildings, electricity is provided by mains connections. For the two sensor stations located in the distribution area within the city, solar panels were placed to provide electricity. Further, specific constructions works were carried out to re-route water through the monitoring station using Pressure Regulation Valves (PRV) to eliminate the need of pumps reducing the energy footprint.

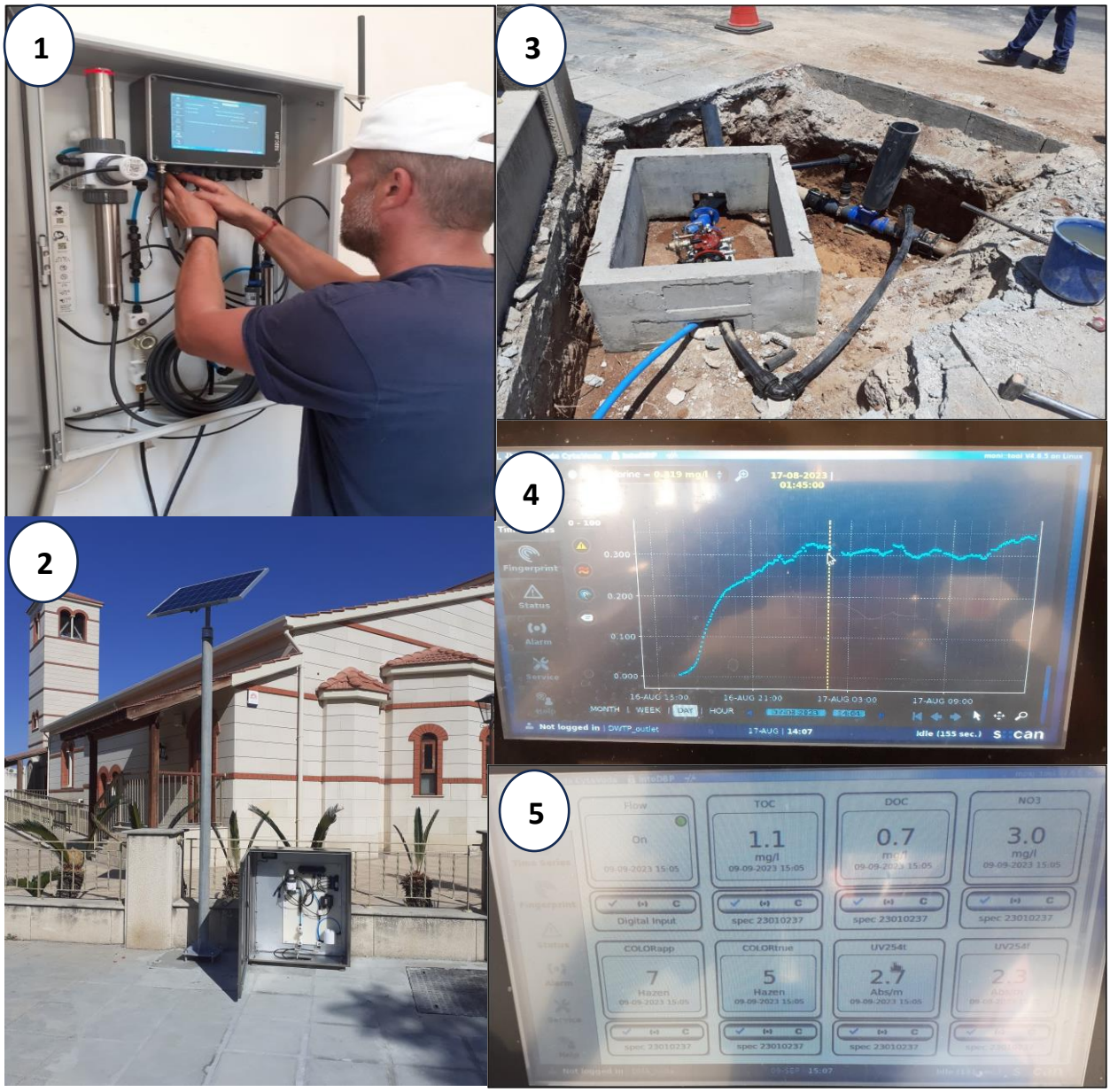


Figure 8: Different works executed for the deployment of the monitoring system: 1) Installation of the sensors and connection with the terminal, 2) Installation of a solar panel to provide electricity for the operation of the sensor station and the transmission of data, 3) construction works and adjustment of pipes for the installation of the sensors, 4) the terminal screen showing time series data, 5) the terminal screen showing the measured parameters.

5.0 Information Communication Technology - Architecture plan

5.1 Data acquisition system plan and ICT diagram

After a terminal read the measurement data from the sensor at the specified measurement interval each data point (measurement data, validated data and meta data) is transmitted by the terminal over a 4G connection, operated with local SIM cards, to a centralized server. The server is located in Germany and operated by s::can. Each terminal has a unique login to the SFTP server used to upload data. The data itself consists of a CSV file generated for each upload and if necessary, a meta data file in the JSON format is transferred alongside this CSV file.

Each SFTP account shares the same folder structure on the server described in Table 2.

Folder name	Purpose
Incoming	Primary folder for data upload from the terminal
Processed	Successfully processed CSV files from the Incoming folder are moved here
Postponed	CSV files from in Incoming folder that could not be processed giving a Python Warning are moved here
Storage	Every Sunday the content of the Processed, Postponed and Failed folders is compressed using TAR and moved to the Storage folder. Then the data is moved to Garbage and kept there for one more week.
Garbage	Garbage folder that is emptied every Sunday before the compression script runs to fill the Storage folder
Failed	CSV files that produced a Python Error are moved here
Logs	Storage for log files

Table 2 of the different folders on the intodbp.s-can.at data server

Files uploaded to the "Incoming" folder undergo analysis by a Python script and are transformed into a standardized data frame, irrespective of the terminal type used. This data frame includes a unique data stream ID derived from the terminal serial number and the transferred parameter value. The data stream ID is stored alongside the measurement data in a Postgres SQL database table, facilitating later retrieval for display or analysis purposes. These CSV files are moved after processing to folders depending on their status (see Table 2). JSON meta data files are kept in the “Incoming” folder as they are needed for the processing of the next uploaded CSV file and only moved if a new JSON meta data file is uploaded due to changes on e.g. the setup of the terminal.

Every successfully processed data point from the CSV file is stored in both the database and the compressed file within the "Storage" folder. Data that finds its way into the "Failed" or "Postponed" folder is exclusively stored in the "Storage" folder. These datasets can be manually imported into the database once the underlying

issues causing warnings or errors have been addressed. A daily backup of the server is performed directly at the hosting company and retained for a duration of 7 days.

Figure 9 provides a summary of how data is managed on the intdbp.s-can.at server. Refer to the next section for an explanation of the user-facing graphical interface.

Additionally, to the automatic data transfer each terminal connects to an OpenVPN server also hosted by s::can in Germany. These VPN connections enable full remote access to all terminals giving the operate the same possibilities as if operating the terminal locally. This allows for remote support of technical issues directly from Vienna without the need to travel to Cyprus or changing the measurement settings remotely if need.

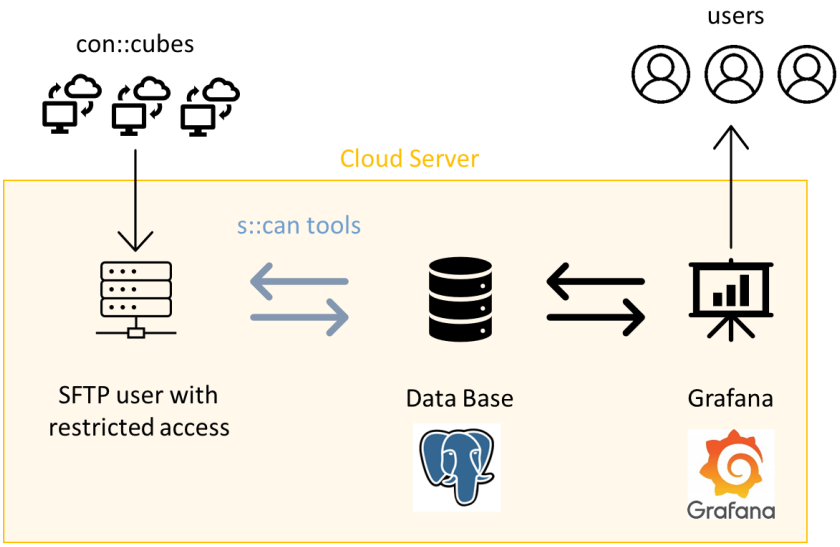


Figure 9: Data management plan as developed by SCAN

5.2 Data visualization and retrieval – User interface

Data stored in the intodbp.s-can.at server's database is made accessible to end users through an open-source web application called Grafana[2]. Grafana enables the creation of tailored dashboards featuring a diverse range of visual representations for measurement data, such as timeseries graphs, gauge views, bar graphs, and more. Dashboards are thoughtfully organized into folders for efficient grouping.

Access to these dashboards is safeguarded by mandatory user logins, each equipped with specific access rights assigned either at the folder or dashboard level. This ensures the fulfilment of individual data privacy requirements.

Data graphs are constructed by crafting specific SQL statements that utilize the data stream ID to access individual data series (see Figure 10, Figure 11, Figure 12). The measurement units, as defined by the terminals, are accessed alongside the corresponding measurement data. The con::cube terminals possess a unique feature with respect to measurement data timeseries: they store two data series - one named

[2] https://grafana.com/

"measured value" and another named "clean value." The distinction lies in the fact that the "measured value" represents the direct output from the sensor, while the "clean value" undergoes analysis and potential modification by a data validation software suite running directly on the terminal.

This software suite examines the noise level and signal level of the sensor, flagging data with warnings in case of deviations from the norm. Additionally, it eliminates outliers by applying a moving average to the "measured data," assuming that changes in water are typically gradual and abrupt data spikes are likely artificial.

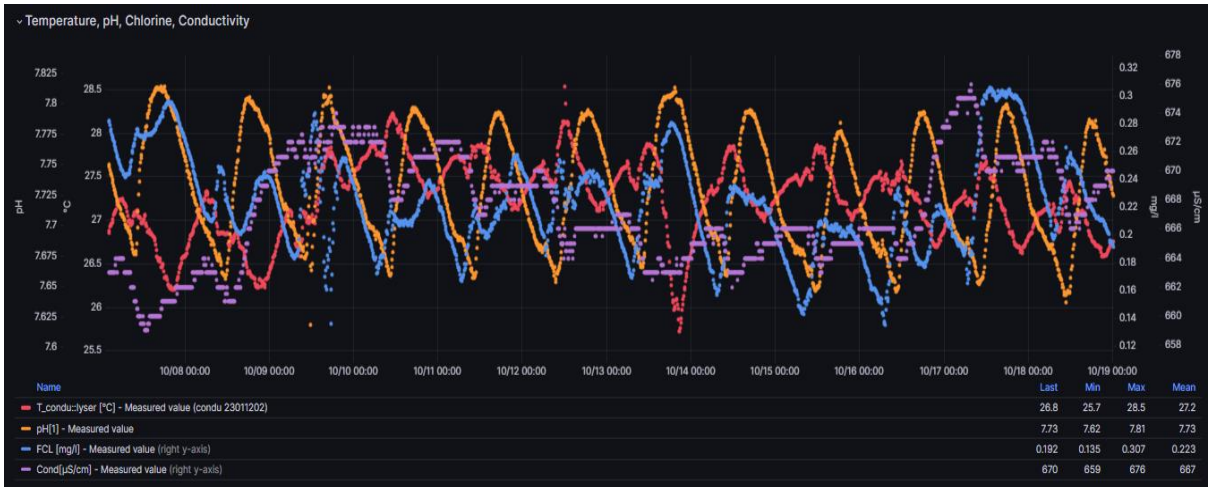


Figure 10: Set of physicochemical data (Temperature, pH, Chlorine, Conductivity) as shown through the Grafana interface

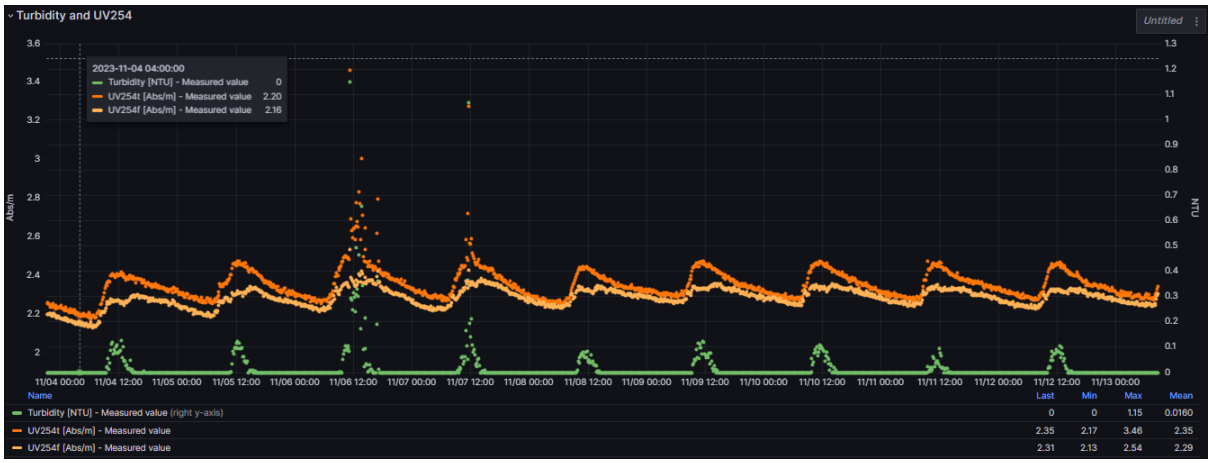


Figure 11: Set of turbidity and UV-Vis data as shown through the Grafana interface

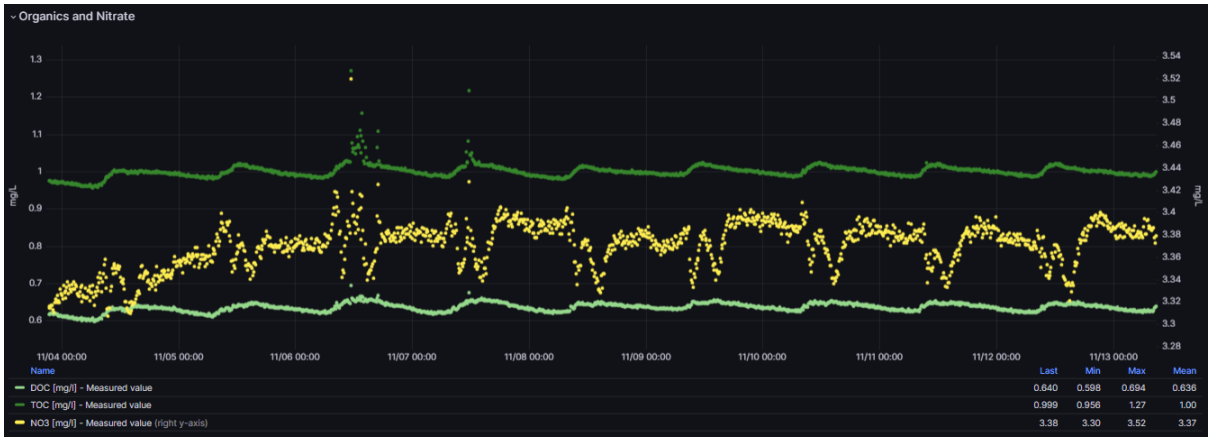


Figure 12: Set of organic and nitrate related data as shown through the Grafana interface

Users have the flexibility to choose their desired time frame through a drop-down menu or directly zoom in on the graphs. The displayed data can be effortlessly exported as a CSV from the menu (refer to Figure 13). However, a notable drawback exists with this export feature: Grafana automatically aggregates data when visualizing large time frames, exporting exactly what is shown to the user. Consequently, obtaining a download of all measured data is not straightforward.

To address this limitation, discussions are currently underway (as of the end of February 2024) regarding new automated methods for data sharing. Presently, accessing time series data across the entire measurement period necessitates either reading the backup files directly stored on the server or exporting specific data streams from the SQL server.

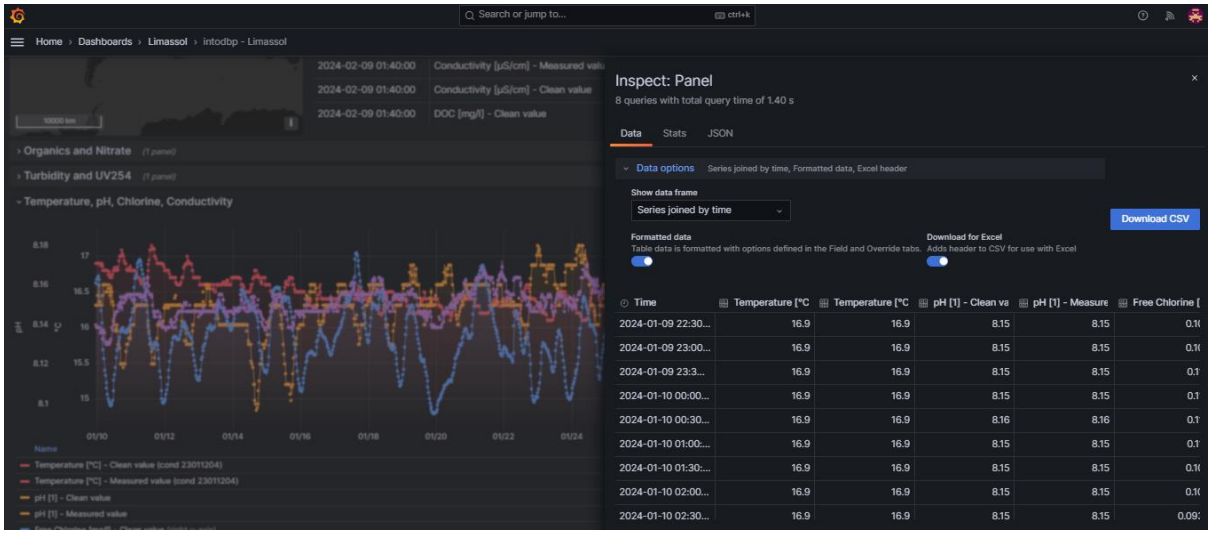


Figure 13: Data retrieval for a selected period through the Grafana interface

Figure 14 and Figure 15 present a sample of data analysed per parameter and per location, respectively. These data were initially exported through the Grafana interface as CSV files, loaded and analysed in MATLAB before integrated with the proposed tool that is currently developed as part of task 3.3.

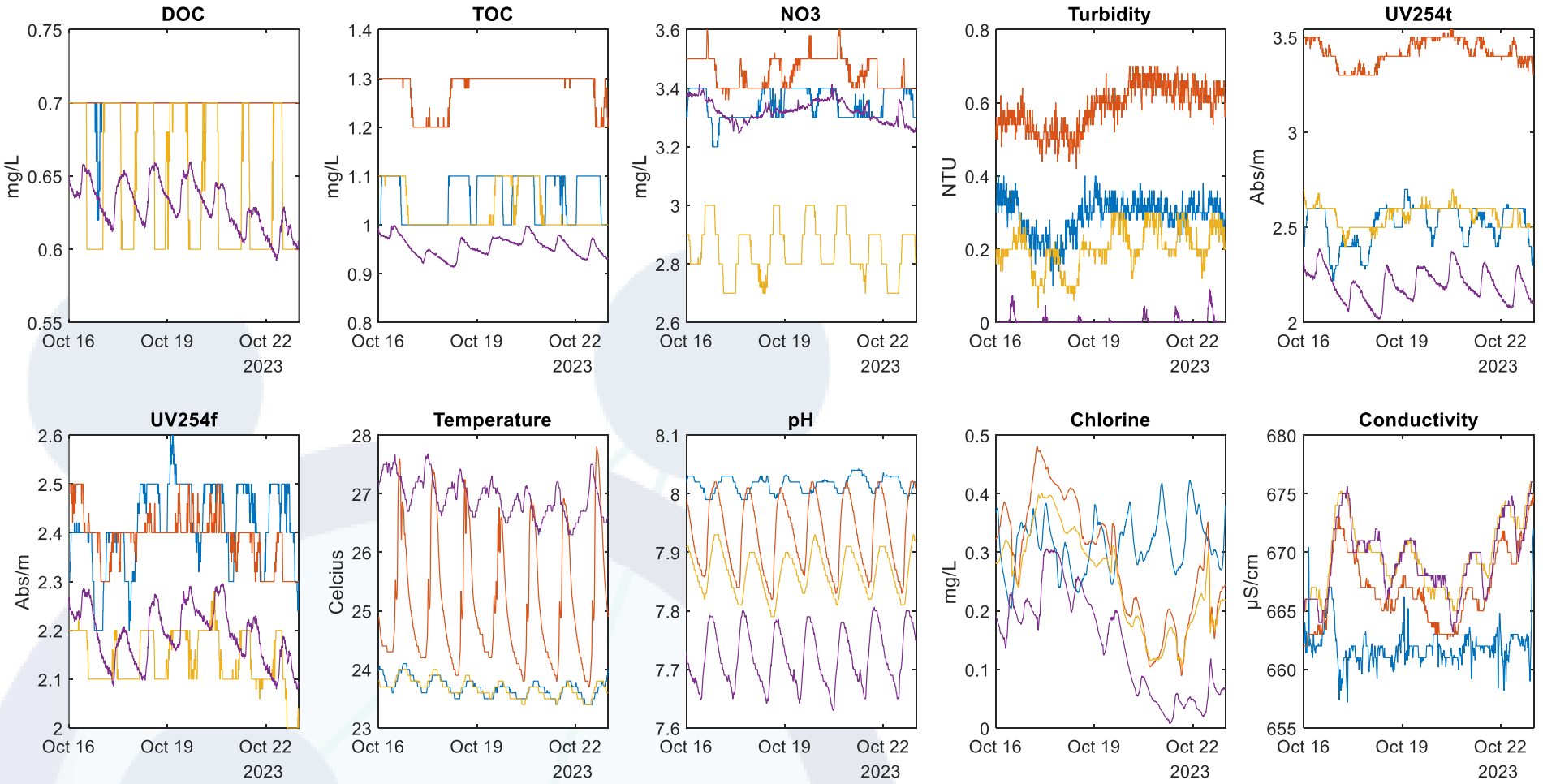


Figure 14: Set of data retrieved between 16 -22 October filtered per parameter

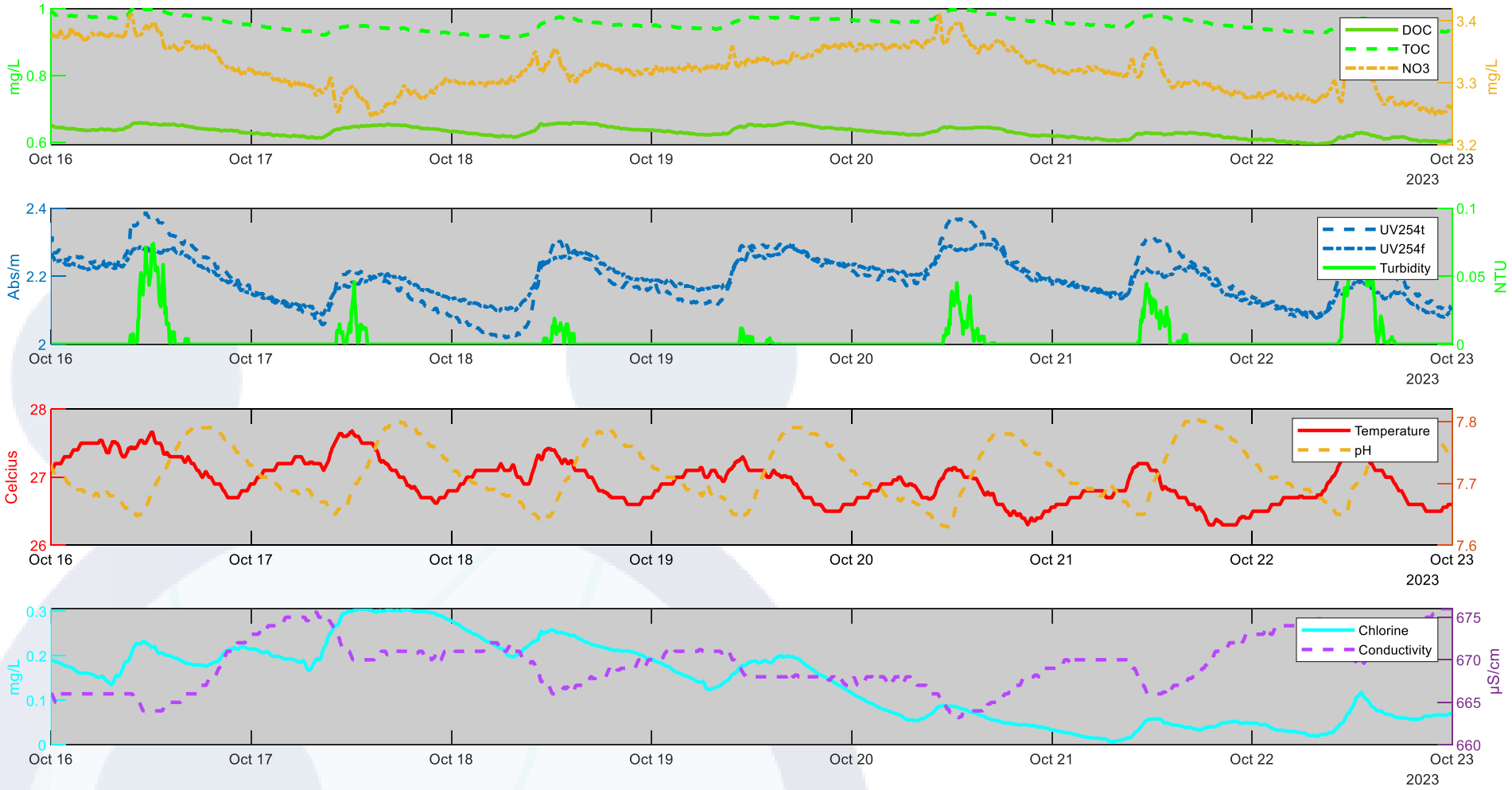


Figure 15: Set of data retrieved between 16 -22 October filtered per location (data from the distribution location)

6.0 Integrated monitoring plan

6.1 Connection with tasks T3.3 and T3.4

In this section we provide a brief overview of how the developed monitoring plan that includes the data acquisition and transmission system will be integrated into a holistic strategy for modelling, monitoring, and controlling the formation of DBPs within the selected case study.

A draft workflow diagram that will be completed within the tasks 3.3 and tasks 3.4 is presented in (Figure 16) and will consist of the following subsystems:

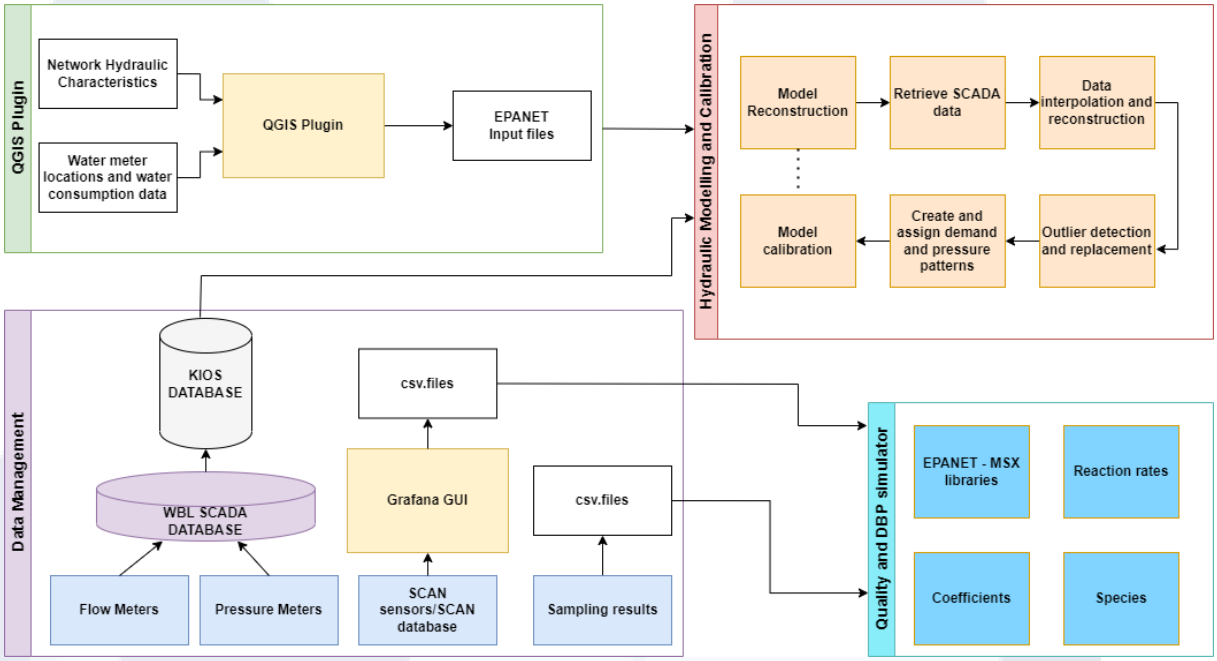


Figure 16: Flow diagram between tasks 3.2 and 3.3. The diagram explains how data retrieved from the monitoring system is used for the development of the proposed simulator of T3.3.

Data Management: This subsystem is responsible for the collection of all data retrieved from the case study and is needed for the development of the proposed simulation tool developed in T3.3. Currently, there is no central database, and the data are separated into three categories: s::can sensor quality data, WBL SCADA sensor hydraulic data and data from the sampling campaign. As aforementioned, sensor data from the WBL are transferred to the KIOS database and can be retrieved through an API using either a MATLAB or Python script. These are routed through a Data Collection (Forwarder) that ensures secure transmission by passing the data through a firewall, thereby protecting against unauthorized access. Data can also be exported as csv. files through a Grafana interface. Data from the s::can sensors, currently can only be retrieved as csv. files from a second Grafana interface and data from the sampling procedure are provided as xlsx. files by NIREAS team who is conducting the samplings. A specific MATLAB script was developed to bring all data in the same form and to integrate them in one dataset. Through this script, the user can request data

for specific time periods and use them in the hydraulic and water quality simulation processes.

QGIS tool and plugin: This subsystem considers the use of the QGIS plugin, "ImportEpanetInpFiles", that was developed by KIOS researchers to facilitate the process of developing hydraulic models (https://github.com/KIOS-Research/ImportEpanetInpFiles). Most of the available information regarding the water distribution network of the selected case study, and in general as well, is provided by water utilities in the form of CAD or GIS files. A preliminary work is initially executed to create accurate GIS models of the network utilizing these files which include the elevation, the pipes' length, diameter, roughness and coordinates, the tanks and reservoirs design and capacity, the valves type and settings, and the consumers locations and consumption data. Following, the suggested QGIS plugin is utilized to convert the GIS models into EPANET hydraulic models.

Toolkits: The retrieved data and the developed hydraulic models will serve as inputs for the following two subsystems responsible for the hydraulic calibration and the modelling of the DBPs formation. To achieve this, the "EPANET-MATLAB Toolkit (EMT)[3]" or the "EPyT: An EPANET-Python Toolkit for Smart Water Network Simulations [4]" will be utilized. Both toolkits were initially created by the KIOS Research and Innovation Center of Excellence at the University of Cyprus and are open-source software designed to function within the MATLAB and Python environments, respectively. They serve as a programming interface for EPANET, a hydraulic and quality modelling software developed by the US EPA. The two toolkits serve as a unified programming framework for research and development within the expanding domain of smart water networks.

Although the other two subsystems, which focus on hydraulic modelling and calibration and the simulation of DBPs formation within the case study, are part of T3.3, a brief overview is presented below:

Hydraulic Modelling and Calibration: This module communicates with the QGIS and the data management subsystems to retrieve all the necessary data for the development of an accurate hydraulic model. The process includes a module for data validation, data interpolation and reconstruction and outlier detection and replacement. Realistic demand and pressure patterns are extracted and utilized for the model calibration process.

DBPs modelling tool: This tool, which is the main output of WP3, will integrate the calibrated hydraulic model and SCAN sensor data with a series of water quality dynamics, residing in EPANET – MSX files, aiming to simulate the chemical reactions between various parameters and predict the formation of DBPs.

Network segmentation module: Creating hydraulic models utilizing GIS data can lead to complex and detailed models which can be computationally expensive when analysed and used in other applications and algorithms. Thus, a network segmentation (skeletonization) process is essential to simplify the model and reduce the simulation

[3] https://github.com/OpenWaterAnalytics/EPANET-Matlab-Toolkit
[4] https://github.com/OpenWaterAnalytics/EPyT

time. The skeletonization algorithm, developed by KIOS, comprise of a set of criteria which must be taken into consideration to achieve a high hydraulic accuracy compared to the original network. In addition, these criteria can serve as selective parameters for a customized *skeletonization*. The sets of criteria include:

- **Set of criteria 1:** Distribution of demands to the end nodes of the new pipes based on four strategies, a) proportional to the distance from the removed junction, b) proportional to the already assigned demand to the end nodes, c) proportional to the elevation of the end nodes of the new pipes and d) equally distributed.
- **Set of criteria 2:** Same pipe characteristics (diameter, loss coefficient, roughness)
- **Set of criteria 3:** Critical elevation (elevation difference between nodes of the merged pipe)
- **Set of criteria 4:** Junctions parameters including the elevation, demand and type (pipe junction, consumption node, fire hydrant, air valve).
- **Set of criteria 5:** Pipe parameters including the status, length, diameter and roughness.
- **Set of criteria 6:** Distance criteria

intoDBP will create innovative tools and strategies to improve water quality management for safe human use and a healthy environment. It focuses on catchment protection and forecasting, transformative drinking water treatment, and real-time monitoring to combat the effects of climate and global change. In particular, intoDBP focuses on pollution and risks related to disinfection byproducts (DBPs). By developing and applying advanced, integrated, and cost-effective sensors and analytical methods, intoDBP will expand knowledge on water quality and DBP precursors to better understand its formation and human exposure in Europe. intoDBP monitoring results will feed into numerical forecasting tools to predict source water changes and formulate climate change adaptation pathways at catchment and treatment scale. intoDBP also develops transformative options for advanced and cost-effective upgrade of water treatment and disinfection.

In the intoDBP consortium researchers, small and large enterprises, communication experts and public services join forces to generate interdisciplinary solutions, that will generate a renewed perspective of drinking water surveillance, support decision-making and governance, and increase system resilience. intoDBP will implement and validate its cross-cutting products in four complementary case studies from three European countries where compliance with DBP regulation currently is an acknowledged challenge. The direct and visible positive impact of intoDBP in the case studies will foster rapid product adoption at a European and global scale, thus strengthening Europe's position and role in the global water market.

Reaching out beyond the water sector itself, intoDBP will directly engage society through surveys to analyze exposure to DBPs, collect data about catchment protection initiatives, create awareness and promote sustainable consumer behavior such as reducing bottled water consumption.

Funded by
the European Union