EARLY WARNING AND IMPROVED DECISION SUPPORT FOR HEALTH PROTECTION IN WATER REUSE AND BATHING WATER

Final version

digital-water.city
digitalwater_eu

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Deliverable N° D1.3	Early warning and improved decision support for health protection in water reuse and bathing water
Related Work Package	WP1
Deliverable lead	ISS
Author(s)	ISS - UNIVPM – KWB – SIAAP
Contact for queries	Susanna Murtas (ISS)
Grant Agreement Number	n° 820954
Instrument	HORIZON 2020
Start date of the project	01 June 2019
Duration of the project	42 months
Website	www.digital-water.city
License	
Abstract	This document describes the implementation of early warning and improved decision support systems for health protection in bathing water and water reuse for agricultural irrigation. Particularly, the document reports description of the planning and design phase, the rationale for the implementation choices, and outcomes of the demonstration.

Dissemination level of the document

X	PU	Public
	PP	Restricted to other programme participants
	RE	Restricted to a group specified by the consortium
	CO	Confidential, only for members of the consortium

Versioning and contribution history

Version*	Date	Modified by	Modification reasons
1	28/02/2022	UNIVPM	Proposed full structure
2	10/03/2022	ISS	Revised full structure
3	15/04/2022	UNIVPM	Add contents to D1.2 (chapter 3 – 3.2.5.1 QMRA)
4	15/04/2022	ISS	Add contents to D1.2 (chapter 3 – 3.2.2.1 Legislative boundaries, Preliminary studies, 3.2.2.3The choice of water quality class at Peschiera Borromeo WWTP)
5	20/04/2022	SIAAP	Add contents (chapter 2).
6	26/04/2022	UNIMI	Add contents (paragraph 3.2.1.3 Agricultural practices in the surrounding areas)
7	03/05/2022	ISS, UNIVPM, CAP, UNIMI	Add contents (Annex A – Risk matrix)
8	05/05/2022	UNIVPM	Add contents (chapter 3 – Digital architecture of EWS)
9	05/05/2022	CAP	Add contents (3.2.1.1 Treatment trains at Peschiera-Borromeo WWTP, 3.2.3.2 Choice of treatment Line 2)
10	06/05/2022	ISS also supported by G. La Rosa, L. Bonadonna, R. Brianceso and A. Coccia	Add contents (Pathogens in raw and treated wastewater)
11	07/05/2022	ISS also supported by F. Simonetti and M. Cerroni	Add contents (Emerging Contaminants in raw and treated wastewater, Statistical analysis of collected data from laboratory monitoring program, Risks prioritization for the SSP at Peschiera Borromeo WWTP, 3.2.5.2 QCRA)
12	10/05/2022	ISS	Add contents (paragraphs 3.2.1, 3.2.2, 3.2.3, 3.2.4, 3.2.5.2- FMEA model and PCA)
13	19/05/2022	UNIVPM	Add contents (Executive summary, conclusion, Annex B – python codes). Integration of partner contributions
14	20/05/2022	UNIVPM	Final revision and editing
R1	26/05/2022	KWB	Revision by coordinator

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

R2	01/06/2022	ICATALIST	Revision by internal reviewer
S	01/06/2022	KWB-UNIVPM-ISS	Submitted version
V	27/01/2023	UNIVPM-ISS-KWB	Revisions to address the comments of the EAB.
V2	03/03/2023	KWB	Minor updates after reviewer comments

* The version convention of the deliverables is described in the Project Management Handbook (D7.1). D for the draft, R for the draft following an internal review, S for submitted to the EC and V for approved by the EC.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Table of content

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

List of figures

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

List of tables

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Glossary

AFRI: Acute Febrile Respiratory Illness

BOD: Biochemical Oxygen Demand

BTX: Benzene Toluene Xilene

BWD: Bathing Water Directive

CEC: Compound of Emerging Concern

COD: Chemical Oxygen Demand

CSO: Combined Sewer Overflow

CoP: Community of Practice

DAFF: dissolved air flotation and filtration

DALYs: disability-adjusted life years

DO: Dissolved Oxygen

DS: Digital Solution

DST®: Defined Substrate Technology

DWC: Digital Water City

E. coli: *Escherichia coli*

EPA: Environmental Protection Agency

EU: European Regulation

EWS: Early Warning System

FMEA: Failure Mode and Effects Analysis

FIB: Fecal Indicator Bacteria

FID: Free Induction Decay

ICAM: Integrated Coastal Area Management

ISO: International Organization for Standardization

IUWRS: Integrated Urban Wastewater and Reuse System

KPI: Key Performance Indicator

ML: Machine Learning

MP: Microplastics

NOM: Norma Official Mexicana

MDSS: Number of demonstrated modelling and decision support systems

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

OUR: Oxygen Uptake Rate

PLC: programmable logic controller

QMRA: quantitative microbial risk assessment

SCADA: Supervisory Control And Data Acquisition

SO: Specific Objectives

SSP: Sanitation Safety Plan

TSS: Total Suspended Solids

WHO: World Health Organization

WQIP: Water Quality Integrated Platform

WSP: Water Safety Plans

WRP: Water Reclamation Plant

WRRMP: Water Reuse Risk Management Plans

WRSP: Water Reuse Safety Plan

WWTP: Wastewater Treatment Plant

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Executive summary

This report describes the development and implementation of Early Warning Systems (EWSs) to support a decision-making process and to assure health protection in water reuse practices and during recreational activities in bathing water.

Particularly, this document presents: (1) the planning and design phase of the early warning systems, (2) the rationale for the implementation choices, and (3) the outcomes of the demonstrations.

The first section of the document (**section 1**) gives an introduction to risk management within an integrated urban water system.

Section 2 describes the rationale for the implementation of a Machine-learning based Early Warning System (DS2) for bathing water quality in the city of Paris. For the development of this EWS, two different models were used to simulate faecal bacteria indicators (FIB) concentration in the bathing sites: a deterministic model (ProSe), and a statistical model. Particularly, data produces by ProSe model were used as input for the statistical model to predict FIB concentration in bathing water. Produced outputs define the sanitary quality of the bathing site according to the Bathing Water Directive thresholds and help managers to decide whether to open or close a bathing site to the public.

Section 3 describes the rationale, design, and implementation of the EWS for safe water reuse (DS3) at Peschiera-Borromeo Wastewater Treatment Plant (WWTP) in Milan. First, the operational framework for the EWS implementation (i.e., technologies applied to treat wastewater, wastewater quality, agricultural practices, legislative boundaries) was described. Then, a semi-quantitative risk analysis (i.e., risk matrix elaboration by WHO indication) and a quantitative risk analysis was performed to detect relevant hazards. Finally, the digital architecture of the Early Warning System was explained, including the development of machine learning algorithms (i.e., soft sensors) for the prediction of target parameters related to relevant hazards. Predictions by soft sensors were demonstrated and validated using real sensors data and laboratory analyses.

The conclusions are reported in **section 4**.

Key innovative elements presented in the report include: i) the development of machine learning models, which were fed by real probe data integrated by data simulated by deterministic model to predict water quality; ii) the contextualization of the developed digital solutions into the framework of risk assessment and management according to the European Regulation for Water Resue and to the Bathing Water Directive; iii) the development of communication tools to share outputs of the EWS and to support decision making in the framework of risk management.

Note: the preparation of this report has been impacted by the COVID-19 pandemics. Due to work restrictions, the supply of the technical material, the installation of the sensors and

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

the field activities have been stopped or delayed of several months. These facts have impacted the implementation and validation of the designed EWSs for health protection in bathing water and water reuse.

This deliverable represents the final version of the interim version submitted in M18 (D1.2). Compared to the previous version it brings additional input regarding:

- Updates related to the development of Early Warning System for bathing water quality (chapter 2)
- Risk assessment and digital architecture of the EWS for water reuse, including soft-sensors development (chapter 3)

Finally, the conclusions (chapter 4) have been revised and updated.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

1. Risk management framework within the integrated urban water system as the basis of safe water reuse and recreational water safety plan

1.1. Introduction to risk management

Risk management is a crucial task, especially in water-related applications, to ensure health and environmental protection. Risk management consists of a continuously evolving process that allows providing the most appropriate measures and procedures to control and minimise negative outcomes. For risk calculation, all the aspects that potentially may affect human health should be considered to ensure an appropriate level of safety. Risks should be identified and managed proactively in order to minimise the risks to the environment and human or animal health.

The risk management approach is increasingly acquiring more importance, and its development is addressed by many research programs, guidelines and standards (Goodwin et al., 2015). The current European bathing water directive (BWD) (EU 76/160/EEC, 2006) demands the implementation of reliable early warning systems for bathing waters. The 'New Bathing Water Directive' was adopted in 2006 and updates the measures of the 1975 legislation and simplifies its management and surveillance methods. Particularly it provides a more proactive approach to informing the public about water quality using four quality categories for bathing waters — 'poor', 'sufficient', 'good' and 'excellent'.

Even the recent European Regulation 741/2020 on minimum requirements for water reuse (EU Regulation, 2020) defines key elements for risk management for the safe water reuse in the context of integrated water management. Particularly, the new European Regulation 741/2020 highlights the need to consider additional requirements depending on specific site-conditions or situations that necessitate particular attention. It could be required to include heavy metals or compounds of emerging concern (CECs), such as pharmaceuticals or microplastics (MPs) in risk assessment. Increasing attention is paid on substances of emerging concern since their impact on health and environment is not yet clearly defined. Risk management should account for potential risks related to CECs, considering additional quality targets other than regulatory standards, but specific procedures are still missing.

Water reuse guidelines encourage the application of Water Safety Plans (WSP) or Water Reuse Safety Plan (WRSP) for both potable and non-potable water reuse schemes.

Water Safety Plans deserve particular attention since they are internationally recognised and well-established approaches. They provide structured methodologies that can be followed by different water schemes. The framework for WSPs should be applied to an integrated system-wide, considering the entire integrated urban water and reuse system from the catchment to the final destination or use.

For this purpose, a superstructure that covers all the urban wastewater cycle, from catchment to its final use, is defined, in order to represent the framework on which the early warning system will be applied. The superstructure considered is the integrated urban water system, which schematises the wastewater cycle in an urban background and includes catchment,

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

sewer network, wastewater treatment plant and final destination. The description of the integrated system is performed in order to obtain a characterisation of each component, defining boundaries and interconnections. Outputs coming out from a step represent inputs to feed the following one. Moreover, multi-barrier systems can address single-stage failure, considering the progressive risk minimisation along the sequence of treatment stages. Available data, assumptions and modelling influence the level of detail and accuracy achievable. Data collected are used to validate the system description. Historical data can be used to estimate system variability. At the first level of analysis, diagram flow schematisation can be used to perform mass balances in steady state. In a further step, real-time data should be integrated, since their use is preferred to represent the real case-study better, whereas models can run dynamic simulations to get the missing information.

Once the system has been defined, all the hazards and the related hazardous events that potentially occur should be individuated along with its structure. Concerning surface water and wastewater, their characterisation and the definition of the possible hazards are strictly dependent on the conditions present in the catchment area and the sewers network. Nonetheless, the expected impacts vary depending on the specific condition of the final environment or destination.

Operational monitoring defines and evaluates the efficiency of control measures applied, and thus needs to consider regulatory requirements, detection limits and data quality, in order to evaluate the effectiveness of control measures in risk minimisation.

Risk management should consider the efficiency of each treatment technology in a multi-barrier approach. Validation of process log reductions is a crucial aspect for the characterisation of microbial risks, in order to define system performance and removal efficiency against pathogens. Stated procedures and clear requirements for the validation of treatment technologies are not always well defined, also considering the wide variety of treatment technologies that could be applied in a system and the technological progress that is always providing innovative solutions not yet standardised.

Risk analysis also introduces some uncertainties, related to the subjective nature of assignment weights and scores, but also concerning the wide variability range of operational conditions, as well as lack of information or poor quality of the data. Even if the presence of those uncertainties is known, there is no specific guidance on how to manage them in the practical applications. Models and analysis, such as multi-criteria decision analysis, fuzzy and stochastic analysis or Monte Carlo based models can help to deal with uncertainty, but their application requires specific skills. Simpler models, such as Failure Mode and Effects Analysis (FMEA) can also be used to assess water systems. Clear guidance on how to assess uncertainties for decision making is required in risk management, but nowadays in current WSPs, it is still missing.

Risks characterisation and prioritisation are fundamental in the development of decision support tools since they allow the identification of the most urgent interventions and facilitate decision-making. Risk assessment is usually conducted by a WSP team, made up by different stakeholders involved and specific expertise, allowing the identification of hazards and

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

hazardous events, characterisation of related risks, the definition of control measures, verification and development of plan's improvements and integrations. Risk management managers, which are the main responsible for the development of the plans, are usually water utilities or irrigation infrastructure managers, often characterised by an operative but a sectoral approach, but have to face with public authorities, which follow a health-based approach.

Finally, the development of supporting programs, stakeholder engagement and communication activities are demonstrated to benefit for risk management. Periodic reports and surveillance to verify the effectiveness of the plan contribute to the continuous upgrade and improvement of the plan. Operative personnel training, research and emergency procedures and manuals can optimise risk management. Communication and public information can help improving confidence in water system, especially in case of sensible applications such as water reuse.

1.2. Early Warning System design: approaches and relationship between risk management and digitalisation

Early Warning Systems (EWSs) combine risk management with digitalisation to ensure safe practices in an automatized and continuous control. Two different approaches can be followed in the design of an EWS. A risk management planning approach can be followed defining, firstly, the risk assessment with the definition of all the hazards, the health target to achieve and the corresponding control measures, and subsequently, integrate digital solutions to control and minimise risks. Alternatively, starting from available data, EWS can be developed as decision support tools, which elaborate data by predefined algorithms and predictive analyses. In this case, risk analysis is implemented in a second moment, providing thresholds for warnings and alarms.

Water management sector usually relies on treatment processes and removal efficiencies, using a huge number of sensors and meters, alarms and automatic control tools. In recent years, technological progress allowed the digitalisation of the water sector providing new sensors, always more precise and reliable, and tools for decision support. On the other hand, health authorities need to rely only on standardised methodologies and certificated data.

Innovative sensors, such as ALERT technology for bacteria measurement, allow more rapid detection of pathogens, decreasing the response time from the 24-48 hours needed for laboratory analysis to 6-12 hours. Moreover, sampling can be automatised in order to perform periodic measurements, even without the presence of operative personnel.

To analyse microbiological risks, health-based targets could be based on indicator organisms, such as *E. coli*, that are easier to measure than specific classes of bacteria, pathogens and viruses. The use of faecal surrogate indicator organisms, such as *E. coli,* is considered in water safety plans when the measurements of microbiological hazards are challenging or expensive. However, the management of data from alternative surrogate indicators is not always well

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

defined in the WSP guidelines. Moreover, it is not clear how to validate data from innovative sensors that do not use standard methods to measure pathogens concentration.

One of the main lacks on the risk management approach is the absence of a common procedure to treat non-standardised data, such as real-time and on-line measurements from sensors or model simulations results. The main difficulty consists in the validation and in the evaluation of the reliability of non-standard data for their utilisation to calculate risks.

1.3. Early Warning Systems available in the market or tested in the literature

EWSs are generally integrated systems consisting of monitoring tools able to analyse and interpret results in real-time (Grayman et al., 2001; US EPA, 2005). The goal of an EWS is to identify the occurrence of low-probability/high-impact contamination events in real-time to make possible the safeguard of public health. EWSs should provide a fast and accurate system to distinguish between typical operational conditions in wastewater treatment plants (WWTPs) and the occurrence of anomalous events or system malfunctions. EWS tools need to be reliable, with few false positives and negatives, not too expensive, easily maintainable, and easily integrated into network operations (Brussen, 2007).

EWSs are currently employed in monitoring systems for drinking water and freshwater quality. Technological advances in instrument development have produced several reliable on-line/real-time monitoring systems able to detect chemical or biological contaminants, treatment malfunctions rapidly, and to assure an optimal water quality management in water treatment and distribution systems. Their application in WWTPs is increasing in recent years, thanks to several technological signs of progress.

Within the water reuse sector, EWSs are mainly applied to control pathogens contamination. For instance, AquaBio analyser is one of the several advanced monitoring solutions tested under the R3Water European project (R3WATER, 2017). AquaBio analyser can quantify Escherichia coli and total coliform automatically in water using the defined substrate technology (DST®), which uses measurements of fluorescence and absorbance for bacteria quantification. The consortium Costa Brava uses AquaBio at their Water Reclamation Plant (WRP) to monitor Escherichia coli and total coliform in both the raw influent and the final effluent of the plant to optimise the inactivation of pathogens (R3WATER, 2017).

In a completely different scenario, the SWIM-Sustain Water MED project promoted water reuse for agricultural purposes in Tunisia (Bedoui, 2014). Particularly, the Médenine WWTP was equipped with a computerised system, which allows the regular sharing of water quality data with stakeholders as well as the early warning notification via SMS in case of quality problems.

Other existing EWSs for wastewater treatment are used for toxic events detection. Toxic contaminants in the influent to the plant can cause inefficiencies in the activated sludge process leading to a reduced removal of organic carbon and nutrients. Chow and colleagues demonstrated the potential usefulness of a detection system equipped with an on-line UV

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

absorbance spectrophotometer to give early warnings of the anomalous operational status at Whyalla WWTP (Australia) (Chow et al., 2018).The on-line spectrophotometer is used for wastewater characterisation, and the acquired spectral data together with other information like rainfall and temperature are managed via the web to detect anomalies. Similarly, in Korea, it was developed an innovative system based on the measurements of dissolved oxygen (DO) and pH to identify potential nitrification inhibition (Hong et al., 2012).In this case, probes were placed in a laboratory scale oxidation/nitrification tank for screening wastewater characteristics, and to early detect toxicity issues due to chemicals flowing into the aeration basin. If the system detects any potential toxicity effect in the incoming wastewater, a WWTP operator can divert the wastewater into a reservoir tank and prevent the inhibition of the biological process. In another study, Du and colleagues assembled an on-line early-warning system to detect toxic loads from industrial wastewaters (Du et al., 2019). The proposed system relies on the measurement of the Relative Oxygen Uptake Rate. It consists of a wastewater tank, a sludge tank, a filter, an aerator, a water pump, a sludge pump, a batch reactor for DO measurement, a DO probe, and a programmable logic controller (PLC). In another work, microbial fuel cell-based biosensor was applied as an early warning device for real-time and in situ detection of Cr(VI) in industrial wastewaters (Zaho et al., 2018)

In addition, there are examples of EWSs that monitor wastewater characteristics in the sewers network before entering the WWTP. It is the case of Lodz WWTP and Wroclaw WWTP in Poland. The advantage is that the warning event is detected early enough to allow WWTP operators to undertake corrective actions. In Wroclaw WWTP, the EWS aims to identify wastewater toxicity level through the measurements of the Oxygen Uptake Rate (OUR) (Jurga et al., 2017). It is possible using few measurement points placed at selected locations of the sewage collection system. Similarly, Black and colleagues presented results from a pilot-scale study using an early warning system able to detect nitrous oxide gas emitted by nitrifying bacteria naturally present in sewer biofilm (Black et al., 2014). In Lodz, instead, a more complex system was realised (Sakson et al., 2019). In this case, the EWS manages different data coming from an existing pluviometry system, an existing flowrate measurements system installed in sewers and placed close to eighteen combined sewer overflows, and from four new stations for monitoring the wastewater quality in sewers. In each of these stations were installed on-line sensors measuring pH, conductivity, organic substances, ammonium nitrogen, suspended solids/turbidity, chlorides, BTX, hydrogen sulphide. The designed EWS aimed to provide quantitative and qualitative information on the wastewater to treat at the receiving WWTP. It allows undertaking corrective actions in advance, and to avoid malfunctions of biological processes.

EWS systems have been also proposed for monitoring bathing water quality providing a surveillance tool for recreational activities. In this context, an interesting EWS was developed in Denmark, where an integrated real-time control and warning system improved the hygienic water quality of the surface waters receiving combined sewer overflows in the city of Aarhus The EWS was operated to reduce the frequency of combined sewer overflows (CSO) (German Water Partnership, n.d.).

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Furthermore, the current European bathing water directive (BWD) (Bathing Water Directive, 2006) demands the implementation of reliable early warning systems for bathing waters. To address this issue, Seis and colleagues proposed an EWS based on multivariate regression modelling, which takes into account the probabilistic character of the European bathing water legislation for both alert levels and model validation criteria (W Seis et al., 2018). The system was implemented using information and data collected at a river-bathing site in Berlin. Precipitation, river flowrates and wastewater discharges were used as key explanatory variables to construct the model. The outputs of this latter study represented the starting point for the development of an Early Warning System for Bathing Wating Monitoring in the city of Paris in DWC project. Table 1 summarizes all the studies and prototype systems realized to provide an EWS tool for wastewater reuse and bathing water monitoring

Table 1: Technological/solution benchmark of proposed EWS for wastewater reuse and bathing water monitoring

Proposed EWS	Project/Reference	Monitored parameters	Application	Goals and comments
AquaBio analyzer	R3Water European project	E. Coli and total coliform	Wastewater reuse	The device provides fast measurements of pathogen concentration in wastewater effluents. The EWS is limited to the measurement of microbiological parameters
Sensors for water quality monitoring and computerised system for sharing information	SWIM-Sustain Water MED project	Typical water quality parameters	Wastewater reuse	Notification via SMS or via internet when non-compliance of wastewater quality occurs. The EWS is restricted to the measurements of conventional wastewater parameters in the effluent
Installation of sensors for monitoring parameters related to biomass activity	Literature studies based on laboratory experiments (Hong et al., 2012; Jurga et al., 2017)	Oxygen Uptake Rate (OUR), pH, dissolved oxygen	Wastewater reuse	Detection of industrial discharges and toxic contamination events. The application is restricted to laboratory studies
Sewer network monitoring systems	Literature studies (Sakson et al., 2019; Chow et al., 2018)	Water quality parameters measured by sensors installed in the sewer network or at the inlet of a WWTP	Wastewater reuse	Research studies have conceptualized/investigated the use of sensors and modelling approaches to detect anomalous load of wastewater entering the WWTP. However, these systems have never been realized at pilot/full-scale
EWS for CSOs control	Case study in the city of Aarhus (Denmark)	Modelling of CSO based on hydrologic and hydraulic parameters	Bathing water	The models aim to predict and reduce CSOs events that affect bathing water quality

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

2. DS2: Machine-learning based Early Warning System for bathing water quality

2.1. Planning and design: description of the solution

The Early Warning System (EWS) for bathing water is, as shown in Figure 1, a system that includes three tools that will target different audiences.

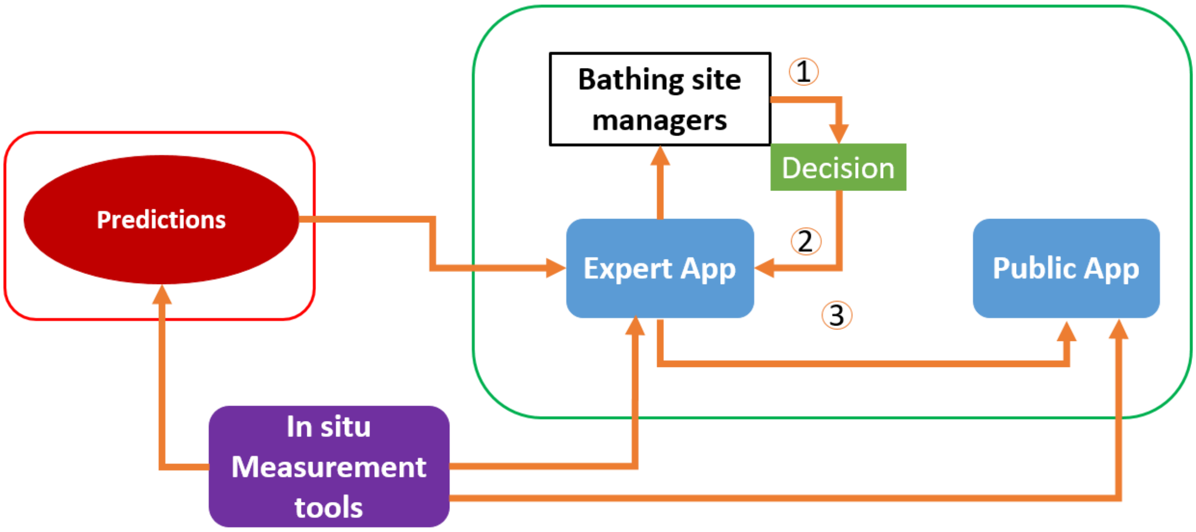


Figure 1 General concept of the Early Warning System

The general concept of the EWS is that the **prediction tool** will simulate the bathing water quality (*E. coli* concentration) including uncertainties at a specific site. This prediction along with additional technical information will be sent to the **"expert" app** that targets the bathing site managers. The expert app is a dashboard which gathers all relevant information to support their decision about opening or closing a bathing site. The **"public" app**, which targets, as its name shows, the public and any stakeholder group interested in bathing water quality (potential bather, boat owners…), will provide the status of the bathing site of their choice as well as additional practical information about the site, bathing safety, technical information…

A Community of Practice (CoP) gathering relevant stakeholders and the future bathing managers has been created in order to determine the requirements for the development of both the "public" and "expert" apps. The aim is to determine the settings and set of information needed (1) by the managers to support informed decision and (2) by the public to foster the use of recreational water in the Paris region.

In parallel, any in situ measurement tools that have been installed in a bathing site will be providing water quality data. These data can be used (1) to validate the predictions of the tool (2) to help the manager to make decision on the bathing site status and (3) to inform the citizens on the quality of the water they want to bath in.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

The originality of the prediction tool for the Paris region resides in the combination of both a statistical and deterministic model. Both models enable to simulate E. coli concentration at the bathing sites.

The motivation of combining the two modeling type where based on multiple reasons. First, the combination allows for using the statistical model for daily predictions. ML based modelling approaches are computationally less demanding and allow for faster prediction. Moreover, online deployment is far easier and requires less resources. Second, the major advantage of the ProSe model is that it allows for accounting for future changes in the sewer network system. While there is plenty historical data available in the Paris case, which generally would allow for training a machine learning model, these data become unusable as new measures for improving water quality are implemented (e.g. UV disinfection at upstream WWTP). This aspect is especially relevant for Paris since under status quo conditions, water quality not suitable for swimming. Thus, the ProSo model is used to identify periods of future good water quality, which the ML model seeks to reproduce. In many cases, where ML models are used for bathing water quality prediction, problems arise due to the limited number of data collected under contamination episodes. In Paris it is currently the opposite, meaning that data availability does not allow for identify periods of good water quality with the required degree of certainty. Third, predictions of the statistical model are site specific, whereas the deterministic model simulates water quality along the river involving a global calibration of the model. Therefore, ProSe allows us to create this kind of data at any place of the river and thus allowing any bathing site to be able to use the prediction tool.

Last, the ProSe model is currently not built to simulate in real time the water quality; it is however, an improvement that is in the work at SIAAP right now. It will take time because it needs a modification of the code of the model in order to be able to process the data in real time.

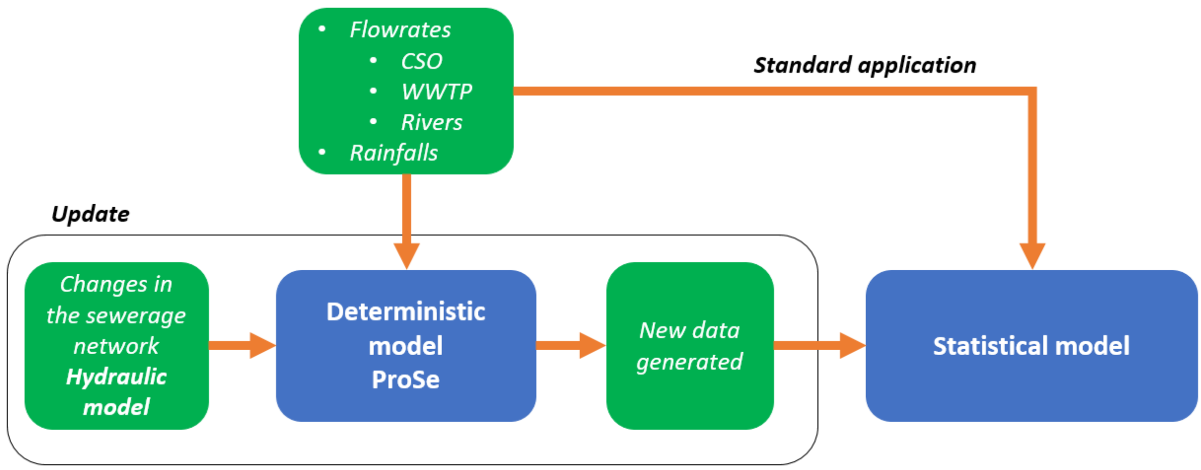


Figure 2 Functioning of the prediction tool

For the daily use of the prediction tool, the statistical model is fed continuously with specific data such as flowrates of CSO, WWTP and rivers as well as rainfall gauges. The model creates

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

a correlation between these data, allowing it to simulate the *E. coli* concentration. This concentration will then be compared to measured data in order to validate the quality of prediction of the model.

The relevance of using both a statistical and a deterministic model lies in the future evolution of the infrastructure. Indeed, while the statistical model allows to predict water quality on a daily basis, it cannot take into account the evolution of the network.This is where the SIAAP's expertise is put forth. A hydraulic model developed by a provider will simulate the different flowrates (CSO and WWTP) taking into account the evolution of the sewerage network. These data will be used as input of the deterministic model ProSe allowing it to produce the water quality at any point of the Marne and Seine rivers.

These new data can be used to re-calibrate the statistical model, which can now take into account the future development of the infrastructure and the impact on the future bathing water quality.

The Early Warning System will be able to monitor bathing water quality at bathing sites, to detect pollution peaks and then alert in case of potential risk for swimmers' health. The FIB concentration and uncertainties predicted by the models will be used to define the sanitary quality of the bathing site considering the Bathing Water Directive thresholds. In parallel with the EWS, regular FIB measurements at bathing sites are necessary to control water quality in real-time. Fluidion's ALERT system could be deployed for this purpose. Moreover, in-situ probes, measuring parameters like turbidity, can be deployed to detect pollution peaks and confirm EWS alerts. These probes can detect punctual pollutions due to involuntary and/or unplanned discharges in rivers that can have a large impact on the water quality of the bathing sites.

2.2. Early Warning System integration in FIWARE

FIWARE will be used to ensure the interoperability of the solution for further replication of the solution. FIWARE will act as a central middleware for data transmission, using FIWARE context brokers and the NGSI-v2 API standard, which makes the setup "powered by FIWARE". Its implementation is realized in close collaboration with Work Package 4. Real-time data will be sent periodically from database (EDEN) and data sensors (Fluidion or others) to the Context Broker. From the Context Broker the data will be used by the statistical model, also referred as open-source software (OSS), for updating prediction as well as the operator to the expert application. The expert and public applications will be connected through FIWARE in order to communicate the predictions of the EWS with the public (Figure 3).

If the user decides to use the model for continuous predictions, the open-source software (OSS, https://github.com/wseis/swim-ai) will be compatible to FIWARE data model standards and FIWARE Orion Context Broker for data transmission. New data, which are needed for updating the predictions will be transferred between the local data providers and the OSS using the Context Broker. The Context Broker will also be used to publish updated predictions generated by the OSS. The 17 processes of continuous prediction are also illustrated in Figure 3, where the OSS is referred to as "statistical model". The application is being tested in Paris

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

and will be connected to the FIWARE interoperability software to predict bathing water quality alongside ProSe (Figure 3).

Considering the fact that the data coming from the database will have a specific format, the purpose of the middleware component (in pink) will be to standardize all of the data in specific data models proposed by FIWARE.

Another important information is that in its current implementation, the Context Broker only stores the last observation, i.e. measurement, from a specific entity (rain gauge, flow measurement). The different apps (e.g. OSS, Expert App) are connected to the context broker using a publish/subscribe implementation. That means that anytime there is a change in a specific entity, e.g. when a new measurement of a rain sensor is published in the context broker, the individual apps receive a notification for that particular change.

The full time series is subsequently stored in app specific databases. For example the OSS uses a SQL database for long term storage, in which the context broker ID is mapped to the ID of the SQL database. For long term storage in the SQL database, existing open standards for data storage based on the "observational data model (ODM)" are used. This architecture ensures FIWARE comparability of the OSS on the one hand, but make the tool more flexible to be transferred other situations where potential other modes of data transfer are used.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

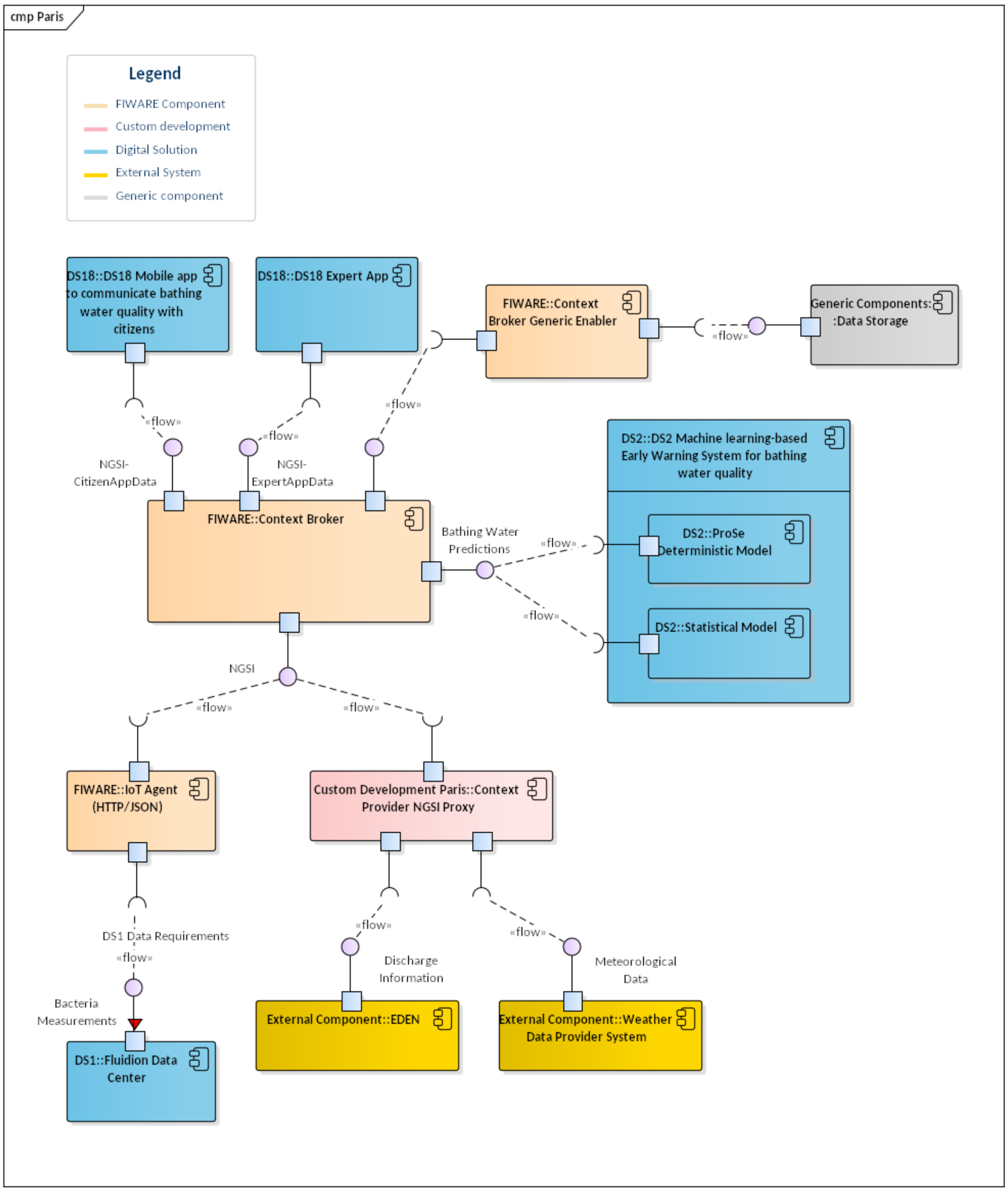


Figure 3: Architecture diagram for the integration of EWS in FIWARE

2.3. WHO Guidelines for safe recreational water environments

Edited in 2003, the WHO Guidelines for Safe Recreational Water Environments describe the state of knowledge regarding the potential impact on human health of the recreational use of coastal and freshwater environments. The main issues addressed were drowning and injury, exposure to cold, heat and sunlight, water quality with particular regard to the exposure to sewage-contaminated water and free-living pathogenic microorganisms, beach sand contamination, exposure to algae and their products, exposure to chemical and physical agents, and dangerous aquatic organisms. The purpose of the Guidelines was to evaluate all

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

monitoring and control measures to protect public health in order to ensure the maximum possible benefit for the bathers and to stimulate the development of international and national procedures (i.e., standards and regulations) to control health risks from hazards that may be encountered in recreational water environments. The application and the implementation of the Guidelines should take into account social, cultural, environmental and economic characteristics, as well as knowledge of routes of exposure, the nature and severity of hazards, and the effectiveness of control measures. Moreover, WHO recommends the conversion and adaptation of the Guidelines “into locally appropriate and applicable standards to ensure a safe, healthy and aesthetically pleasing environment”. The conversion of the guidelines into regulations adapted to local circumstances needs the consideration of several arguments for each type of hazard. In the case of recreational water, the main requirements for a quality classification due to faecal pollution should include: i) the establishment of a water quality classification system; ii) the obligation upon the national or appropriate regulatory authorities to maintain a list of all recognised recreational water areas in a publicly accessible location; iii) the definition of responsibility for establishing a plan for recreational water safety management and its implementation; iv) independent surveillance and provision of information to the public; v) the obligation to act, including the requirement to immediately consult with the public health authority, and inform the public as appropriate on detection of conditions potentially hazardous to health; vi) a general requirement to ensure the safest achievable recreational water use conditions. The process leading to the application and adaptation of local standards and guidelines needs the involvement of multiple stakeholders. To achieve this objective, the establishment of a coordinated management system for recreational marine and freshwater areas based on an Integrated Coastal Area Management (ICAM) is crucial. This involves a comprehensive assessment, the setting of objectives, the planning and management of coastal systems and resources. It should also take into account traditional, cultural and historical perspectives and conflicting interests and uses. In an ICAM program, the exact package of management options to reduce or eliminate health hazards and risks related to recreational water uses will be driven by the nature, frequency and severity of the public health impacts.

2.4. Case Studies of Risk Management Plan for bathing waters

Exposure to contaminated recreational waters can lead to diseases, especially for susceptible people with reduced immune function. Water quality highly depends on the anthropic pressure (especially in heavily populated areas located in the proximity of industries and agricultural activities). Particularly, wastewater discharges can vehicle contaminants (e.g., pathogens and chemicals) increasing the risk for humans to contract water-related diseases. Even though several pieces of evidence have shown some correlations between adverse effects on human health and poor quality of recreational water, there are still difficulties to identify the actual cause (pollution) - effect (disease) relationships. Most of the scientific investigations in this field have been focused on infections associated with recreational waters resulting in minor, self-limiting symptoms. Indeed, the attribution of severe illness to recreational water exposure is challenging to establish due to the great number of

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

transmission routes for pathogens. Furthermore, it seems that some microorganisms or their products may be directly or indirectly associated with secondary health issues. Nevertheless, it exists some rare evidence that the risk to get some potentially fatal disease can be related to the quality of recreational water. The first assessments of the incidence of recreational water related diseases were published in the early 1920s in the US Public Health Association. Other epidemiological studies on the relationship between bathing waters and illness were conducted between 1948 and 1950 by the US public health Service. Extensive studies in five years on 43 beaches, which were conducted in the UK from 1959, concluded that there was a "negligible risk to health" of bathing marine water polluted by sewage, even though the investigated beaches were classified as "aesthetically very unsatisfactory". The study stated that a serious risk would only exist if water were so fouled as to be revolting to the senses. The issue remained controversial for many years until the United States Environmental Protection Agency (US EPA) established in 1972 that there was a lack of valid epidemiological data to set guideline standards for recreational waters. The Guidelines for Safe Recreational Water Environments edited by WHO (WHO, 2003) concluded that the most frequent adverse health outcomes from recreational waters are gastroenteritis and acute febrile respiratory illness (AFRI). The guidelines confirm the association between gastrointestinal symptoms, AFRI and indicator-bacteria concentrations in recreational waters. The need for stronger recreational water monitoring programs to reduce the development of related illness has been acknowledged by several institutions worldwide. In 1999, The US EPA introduced the Beach Action Plan (US EPA, 1999) to describe actions to improve and assist in the state, tribal, and local implementation of recreational water monitoring and public notification programs. The primary objective was to enable consistent management of recreational water quality programs to identify the needs and deficiencies of recreational water quality monitoring. The second objective was to improve the knowledge in support to recreational water monitoring programs addressing three broad arguments: (i) Water Quality Indicators Research, including rapid analytical methods to identify risk before exposure occurs; (ii) Modelling and Monitoring Research, a joint estimation between computer models and laboratory tests to generate a reliable determination of health risk; (iii) Exposure and Health Effects Research, to determine pathogen occurrence and indicator relationships associated with wet weather when sewer overflows are combined with discharges of stormwater. The plan was implemented within a workgroup comprising expertise in water monitoring leading to the development of guidance for public health professionals on when, where, and how to set up and conduct an appropriate monitoring program for typical beaches. An example of a risk management plan on inland waters is the "Isar plan", which is a water management action for the restoration of a stretch of the Isar River in Munich, Germany. Main aims of the Isar Plan were directed to the improvement of flood control by increasing the water retention capacity of the river stretch; to protect habitats for wild species, biodiversity and water quality; to the improvement of the water quality of the recreational scope. About this latter point, the plan aimed to find a suitable solution for the growing need for recreational space in a densely populated urban area. Actions to improve the recreational use of the river water consisted of widening the riverbed and developing banks and flat ramps with intermediate pools using natural solutions. The plan was developed within an interdisciplinary working group including the State Office

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

of Water Management Munich, the City of Munich and the "Isar-Allianz" (an alliance of NGOs). The working group examined the flood risk, the need for recreational areas at the riverside and the area's animal and plant worlds and their habitat. Based on their findings, the development goals were defined.

2.5. Local deployment of the solution – Paris case study

2.5.1. Deterministic Model

For the city of Paris, the modelling of the natural environment processes is a relevant research topic of the PIREN-Seine programme (Interdisciplinary Research Programme on water and the Environment of the Seine basin), and the development of ProSe modelling is part of it.

The deterministic model ProSe was developed in the context of the PIREN-Seine programme (Interdisciplinary Research Programme on water and the Environment of the Seine basin) for the study of Seine and Marne rivers in Paris. It includes hydraulic modules, as well as modules for the simulation of Faecal Indicator Bacteria (FIB) (Servais et al., 2007). In fact, ProSe includes different sub-models to simulate several physical and chemical parameters of water in the Seine and Marne rivers in order to have a better understanding of water quality variations related to chronic or accidental pollution. In this context, a module for the simulation of Faecal Indicator Bacteria (FIB) has been implemented in ProSe over the period 2007-2010 (Servais et al., 2010) and serves as a starting point for the elaboration of the EWS for bathing water quality in Paris. Moreover, a wide amount of data is available for Paris area, and it represents a great opportunity for the development and calibration of data driven models. The four major types of data collected from different providers are flowrates; chemical parameters (such as NH4, Dissolved Organic Carbone, Chemical Oxygen Demand…); FIB concentration; and rainfall. The association of a large water quality dataset and a robust modelling approach are key factors to improve the rationality of the decision-making process. ProSe simulates the evolution of water quality along the Seine and Marne Rivers. The river state at each simulated point is defined by physical variables (like water speed, water level…) and biochemical variables, which might be influenced by other factors like temperature. In order to simulate these variables, the model uses different modules such as the hydraulic module or the water quality module describing FIB variation dynamics (Servais et al., 2011). The model allows linking the river's metabolism with the anthropogenic pressure due to urban discharges. Figure 4 presents a flowchart of how ProSe works.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

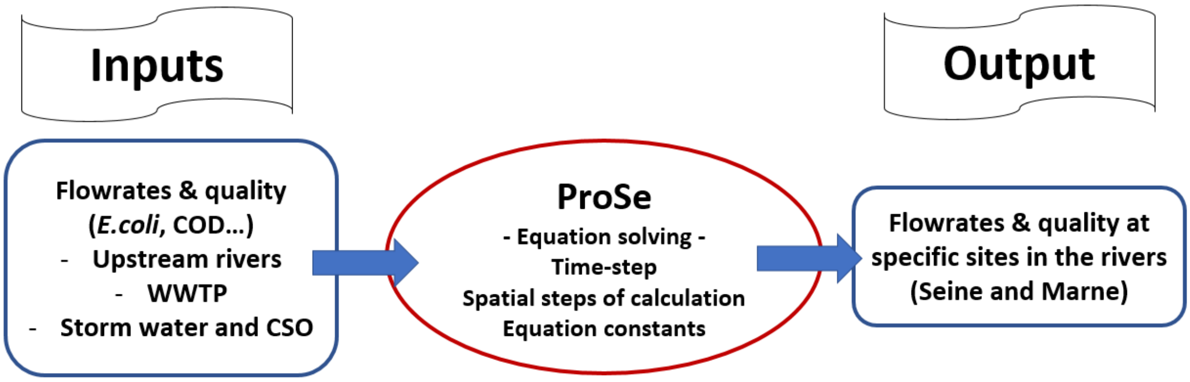


Figure 4 ProSe modelling

The deterministic model can be adapted to simulate infrastructure changes in the sewer network or future scenarios for CSOs and WWTPs discharges. In fact, since it is connected to different sub-models, each of them specialised to investigate a definite aspect from hydrology to bacteria contamination, ProSe includes the representation of the physical system and thus is suitable to evaluate different scenarios, including future designs and infrastructure changes.

On the other hand, the accurate description of the real field conditions requires a high level of complexity, with long calculation time. The process workflow implies first to run simulations from each sub-model, then the simulated intermediate outputs must be fed manually as inputs to the ProSe. Moreover, since the interest was focused on the urban city of Paris, efforts were required to reduce the extension of the simulated area. ProSe provides hourly values of bacteria concentration, but results do not include information about accuracy or sensitivity analysis yet, even if new upgrades are going to be implemented to take into account uncertainties. Among technical and operative issues, one of the main obstacles to the exclusive use of ProSe as EWS is that, due to its structure and the complex mechanism for feeding input data, sensors cannot be directly connected to the model to provide predictions, making ProSe not suitable for real-time simulations. Nevertheless, its strong connection with the physical system and the possibility to perform different scenarios on future configurations, make ProSe a highly valuable tool to obtain a robust dataset of bacteria concentration trends, that could take into account different environmental conditions and system configurations.

Simulations with ProSe require large efforts in terms of time and expertise and thus it could be not suitable for real-time evaluations and EWS. However, its flexibility in representing different site conditions and its ability in providing a robust set of data were valorised to calibrate a statistical model, that is easier to use, time saving and feasible for real-time EWS, but that alone would be not flexible to system changes and has to be trained with a dataset that include as many variabilities as possible to provide reliable results.

2.5.2. Statistical Model

The digital approach taken in Paris is complemented with experiences and methodologies developed in Berlin under the project FLUSSHYGIENE. The approach is based on readily

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

available data, like rainfall, river flow and WWTP discharges and uses probabilistic forecasting for predicting FIB concentrations after heavy rainfall. The approach, moreover, implements risk-based decision making as outlined in WHO guidelines and the EU bathing water directive by extending the probabilistic approach for bathing water quality assessment ("percentile-approach") to short-term decision making.

Water quality at the bathing site might be impaired concurrently by both rain events that occurred just recently and rain events that happened several days ago. The transit time of a contamination from its point of rejection to any bathing site depends on the river flow Q. In order to account for these kinds of variations, different explanatory variables are to be constructed. From WWTP discharges as well as from flow data Q, daily sums (WWTP) and averages (Q) are calculated up to seven days before sampling for each individual day (Q1, Q2, ... Q5 and WWTP1, WWTP2, ... WWTP5). Moreover, variables are constructed which summed/averaged over multiple days before sampling, e.g., Q1-3, WWTP 1-3 as the average/sum over three days before sampling. Cyterski et al. (2012), Herrig et al. (2015), and Seis et al. (2018) have already successfully applied a similar approach. Rain variables are created analogously including a log-transformation. A value of 1 is added before taking the logarithm. The rationale for log-transformation is that while discharges of CSO and storm water may increase FIB concentration by orders of magnitude in the first instance, a further increase in rain and consequently discharge volume will not increase the concentration linearly on a log10 scale. By log-transformation, the effect of higher rain levels is weakened. The sampling day will not be included in the averaging, since precipitation might have started after sampling in the case of historical data, creating artefacts of wet weather conditions, when the sample might actually have been taken under dry weather conditions. Due to the lognormality assumption, given by the BWD, *E. coli* data were log10-transformed. Thereby, fitted models will be able to predict the lognormal distribution (mean and standard deviation) of faecal indicator bacteria and will be in line with current European Bathing Water legislation. The latter uses the percentiles of a lognormal distribution for probabilistic long-term bathing water quality classification. The applied approach translates the thresholds from the European bathing water directive to short-term predictions

In digital-water.city (DWC), this approach is applied into an open-source software that supports decision-making by training and validating probabilistic prediction models based on readily available data. The user will be able to upload paired datasets (i.e., FIB data + time-series data predictor variables). These datasets represent historical data and are used for model calibration (supervised learning). The open-source software (OSS) evaluates multiple machine learning algorithms for automatic variable selection and model validation on the provided dataset. The OSS provide the user with information about the quality of the different learning algorithms. Moreover, it returns historical bathing water quality predictions based on the provided data. These historical predictions allow the user to assess how the model would have predicted bathing water quality in the past and thus, provided information about the practical implication of implementing the model.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

2.6. Data collection

The study period for the development of the EWS covers four years from 2016 to 2019. The mandatory data for Prose simulations are flowrates and quality in WWTP and stormwater overflows. These data are collected at high frequency (5 minutes time step) when available. It represents a high amount of data, as there are around 115 combined sewer overflows or separated stormwater overflows in the simulated area (see figure 5). These data were collected for the main discharges in the Seine and Marne rivers, which included the Wastewater Treatment Plant (WWTP), the rain gauges, and some river water samples located at the potential bathing sites in Paris. Other data of interest are river quality and microbiology analyses. Those last two can be used to compare the simulation with measured data at some key locations in the river. In addition, measurement campaigns using the ALERT System have been conducted to provide FIB data in strategic bathing sites such as the Alma Bridge, which is close to the bathing site of the 2024 Olympic and Paralympic Games. On a wider scale, the routinely measured FIB concentration data from grab samples at several locations in the Ile-de-France region were collected from various partners involved in the water sector. This data collection will also serve as the foundation for the statistical model developed by KWB.

Complementary to the data collection, measurement campaigns have been conducted in order to improve our knowledge of the water quality in urban discharges and rivers. The location of the monitoring sites of 2019 and 2020 is presented in Figure 4.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

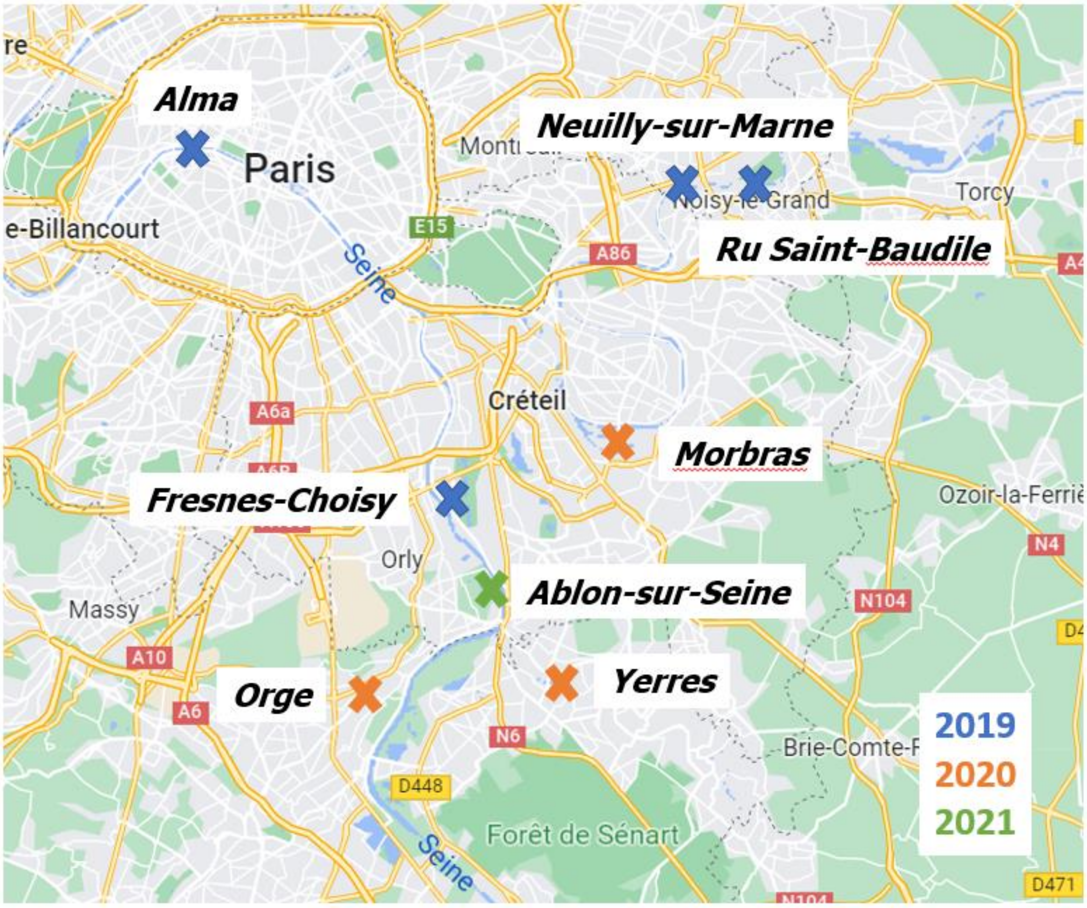


Figure 5 Measurement campaign

The 2019 campaign focused on the Seine and Marne rivers. More specifically on two bathing sites (Alma Bridge and Marne) and two major discharges points (Fresnes-Choisy in the Seine River and Ru Saint Baudile in the Marne River). In addition to a temporal variability campaign on the four sites, which included microbiological and physic-chemical monitoring, a low scale spatial variability campaign (100 to 200 meters long for 12 spaced-apart samples) was conducted at the shores and middle of the rivers at the two bathing sites (Alma Bridge and Marne). This last campaign was especially realized to compare the difference between the shore and the middle of a river as a sampling point. Both the laboratory and Fluidion's ALERT system V1 carried out the analysis of the temporal variability campaign. However, only Fluidion provided drones dedicated to the spatial variability campaign.

The 2020 campaign focused on smaller rivers that are tributaries of the Seine and Marne Rivers. The "Yerres" and "Orge" Rivers were monitored as the Seine's tributaries and the "Morbras" as the Marne's. The objective was the determination of FIB contributions from upstream tributaries during dry weather and rainy weather. As the hydrographs were quite long for these tributaries (from 1 to 3 days), it has been decided to monitor the rainy events over two days. Which represented 12 samples averaged over 4 hours. An external provider installed automatic samplers at each of the sites, triggered by means of a GPRS

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

telecommunication box allowing the control of a relay via SMS. Once these samples delivered to the SIAAP, a certified laboratory realized the *E. coli*, intestinal enterococci and 22 physico-chemical parameters measurements. In addition, the same provider installed in-situ probes near the samplers to monitor conductivity and turbidity.

The 2021 campaign focused on the Seine River upstream of Paris right before the dam of Ablon-sur-Seine. The ALERT System V2 was installed for two whole months (June and July) in order to study the water quality during dry and wet weather. In parallel, a technician from SIAAP did a manual sampling once a week at the same time as the ALERT System V2 in order for the data to be compared.

The summer of 2021 was a peculiar one. Indeed, it was raining a lot so getting water quality data during wet weather was not difficult, the ALERT System was set to analyse a sample every 6 hours on a 48 hours timeline. As for the dry weather, a sample was taken every day at the same time however, considering the amount of rain, it was decided as soon as real dry weather appeared to do a 24 hours sampling every 3 hours.

In addition to that, the weekly sample analysed by the certified laboratory also allowed us to get NH_4 concentration and total organic carbon.

This last campaign allowed us to get additional data from a site upstream of Paris that would be used to help calibrate better both the deterministic and statistical model.

Additional FIB data are being collected every year by different stakeholders during the summer in the Seine and Marne rivers to improve our knowledge of the water quality in wet and dry weather

2.7. Monitoring program and model calibration

2.7.1. Model calibration - PROSE

Current simulations consider a perimeter around Paris with a start in Choisy for the Seine River and Neuilly for the Marne River and an ending at the Suresnes dam, downstream of Paris. The initialization sites will be shifted farther upstream during this project to increase the geographical coverage of the model. ProSe requires input data to start a simulation.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

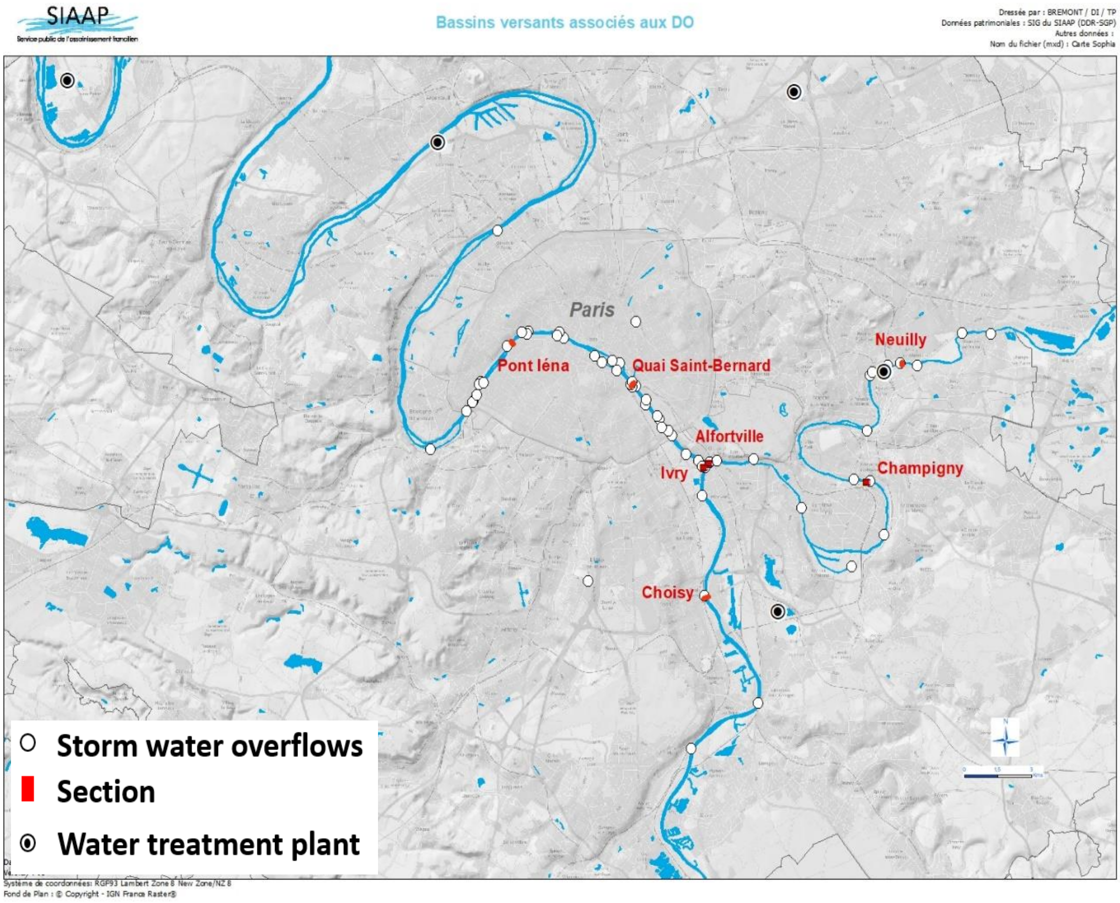


Figure 6: Study zone for the ProSe model

For a better understanding of the study zone, it was divided into sections:

On the Marne River:

- From Neuilly to Champigny
- From Champigny to Alfortville

On the Seine River upstream of Paris

- From Choisy to Ivry

On the Seine River in Paris

- From the confluence to the Quai Saint Bernard
- From the Quai Saint Bernard to the Pont of Iéna.

The main inputs are flowrates and quality at the different upstream rivers. Then, there are the discharges in the rivers with on one side the WWTP, and on the other side, the stormwater overflows that mainly occur during rainy weather. Once these data entered into ProSe, the simulation can start, and the model solves the different equations of each module to

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

determine the flow and quality all along the selected river section. First simulations results describing FIB dynamics during the summer of 2010 showed good agreement between measured and simulated FIB with, however, some notable differences during wet weather events (Poulin et al., 2013). Other actions are being conducted to improve simulation accuracy. It includes a better estimation of the input FIB concentrations in stormwater and combined sewer overflows and increasing the simulation frequency to provide a better representation of short-term pollution events. Finally, it involved the improvement of the modelled phenomena by refining the determination of the variables used for the growth and mortality of FIB in the ProSe model as well as their distribution in the different layers of the river (water, total suspended solids and sediment). The quality of the simulation will be further improved by achieving the on-line integration of data measured in real-time. The generated simulation output will allow the creation of data series to calibrate the statistical model considering different development scenarios of the sewer network and WWTP. Indeed, the deterministic model allows the modification of the study area according to future changes. ProSe allows simulating the evolution of discharges in the near area, while smart integration of deterministic and statistical model leads to an adaptable Early Warning System that is able to consider the ongoing changes of the urban water infrastructure. The first action in progress is to determine the optimal time-step for input data. Indeed, current simulations are done using 24h average values for input data of discharges. This low frequency results in an overly smoothed discharge profile and is not appropriate to assess the variability of FIB concentration in the river during rain events. Therefore, it is necessary to define the optimal time-step by assessing the sensitivity of the model to various time steps.

2.7.2. Model calibration - Statistical model

Machine learning (ML) models need to be calibrated and validated using large datasets. In the field of supervised learning, which is used for the specific application in DWC, paired data of the target variable and suitable predictor variables have to be collected. For this implementation, the concentration of faecal indicator bacteria is used as the target variable. To do so, SIAAP, KWB, and the University of Sorbonne agreed on using an existing data model (Water ML, ODM2) as a template to organize the data exchange between Berlin and Paris. Note, this data exchange refers only to the transfer of historical data used for model calibration. For setting up the real-time data transfer, which is used for regular predictions, the FIWARE Orion Context Broker is used, as described in section 2.2. An overview of the different data sources is given in Figure 7 (Rainfall stations are not shown to allow readability of the map). From the collected data, multiple features were engineered. Subsequently, a variety of ML modelling approaches were fitted and tested to the collected data sets, including:

- Multivariate Bayesian regression modelling
- Penalised regression approaches (Lasso, Ridge)
- Tree-based modelling approaches (Random Forest, Bayesian Additive Regression Trees)

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

As a first approach, model validation will be based on internal cross-validation approaches. In the future, the simulations realized with both models will have to be validated by real on-site measurments.

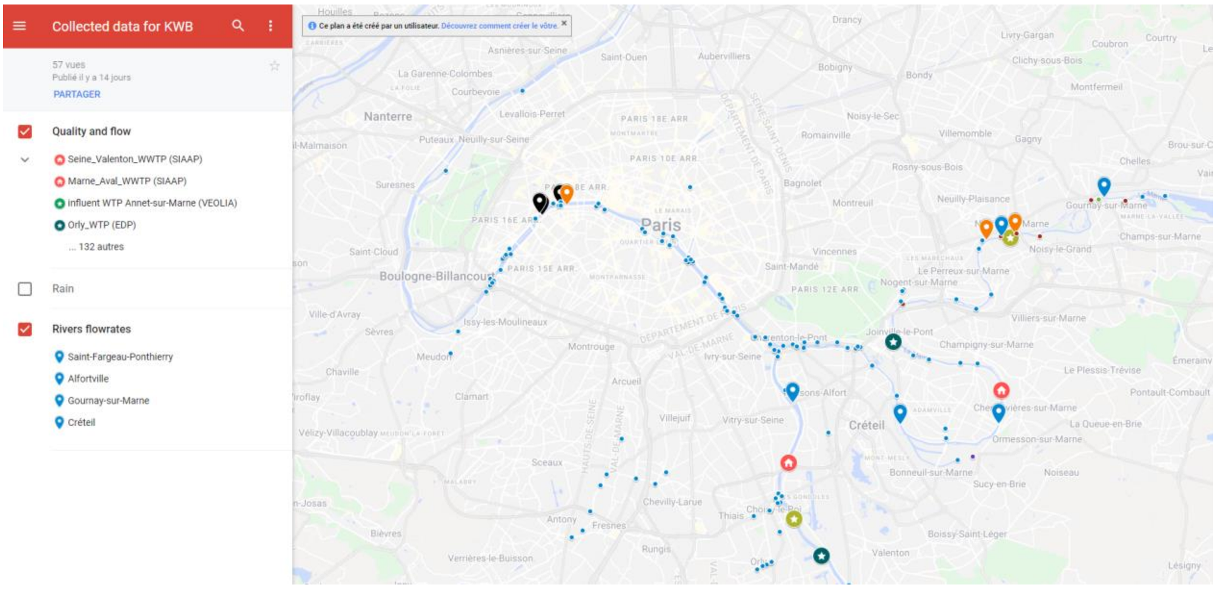


Figure 7 Overview of used predictor variables for calibrating ML models in Paris

2.8. Firsts outcomes and validation results

First tests were carried out using the simulated data provided by ProSe, in order to evaluate the level of agreement between the two models and verify the calibration performances using simulated data.

As part of its sanitation master plan, SIAAP constructed different scenarios putting forth the multiple changes that need to occur on the sewerage network in order to reach the water quality for bathing. Those scenarios include among other things the disinfection of the WWTP, the correction of some connection that led to CSO and the construction of collecting basins. The impact of 3 scenarios (SC4, SC6 and SC7) were studied and simulated using the deterministic model ProSe.

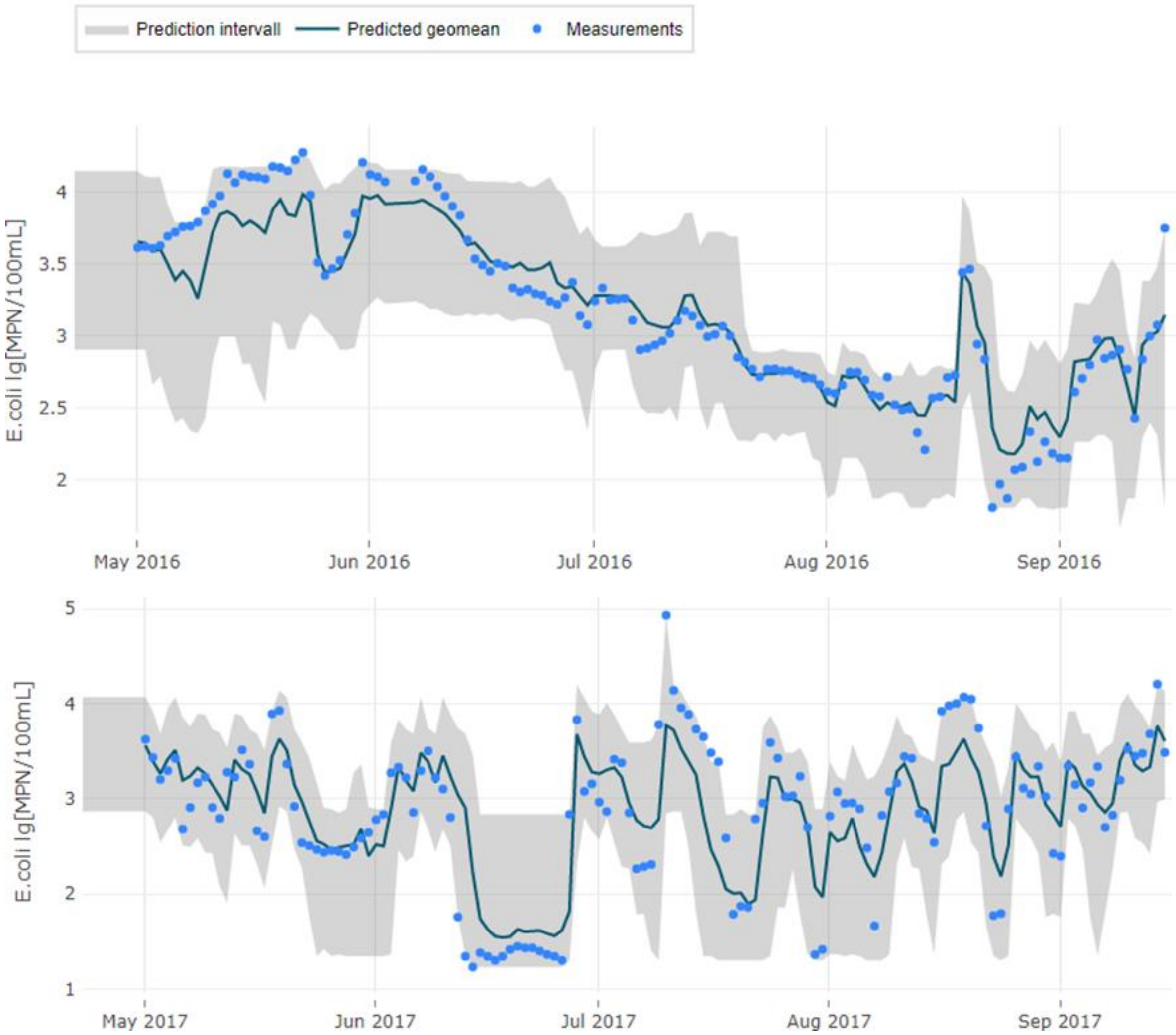
Outputs from statistical model are shown in Figure 8 as median trend (blue line), coupled with the uncertainty range (grey area), while ProSe data (represented as light blue dots) were not provided with their uncertainty.

Results showed in general a good agreement between ProSe and the statistical model. However, some discrepancies can be detected. It has to be noticed that, however, in correspondence of biggest discrepancies, the uncertainty range of the statistical model also increased, meaning that the model was able to understand that in correspondence of those periods its results could have been affected by bigger errors.

The differences between the two models' outputs are focused on specific time periods and could be probably attributed to extra-ordinary maintenance interventions, that could have

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

caused anomalous discharges. It has to be noticed that currently the statistical model is using data from WWTP discharges, river flows and rain flows, since they are data that are more often available and that could be easily collected. However, for the future implementation of the statistical model, CSOs discharges will be included in the analysis, in order to get all the relevant information that could affect the results.



digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

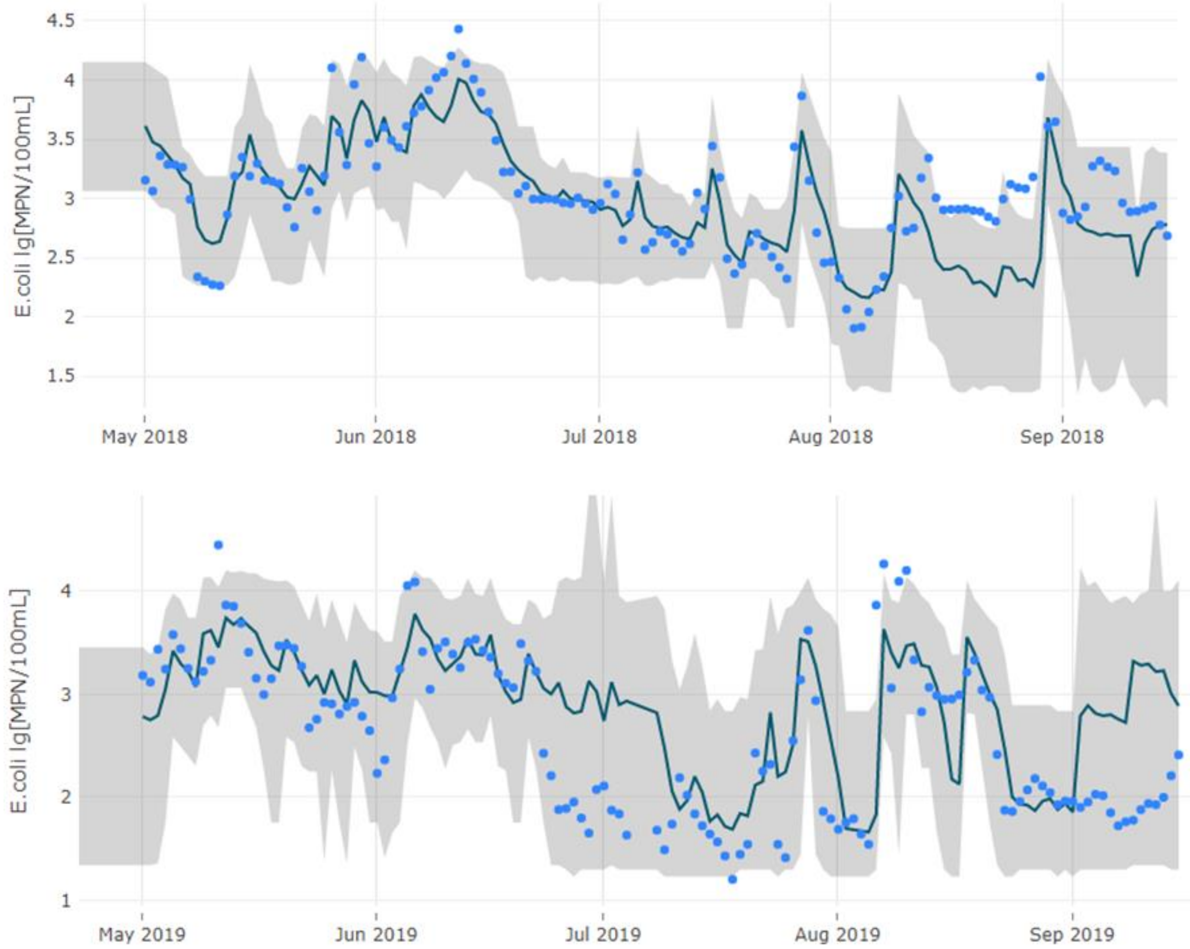


Figure 8: ***E. coli*** **trends for scenario SC7 in summer period of years 2016-2019**

From the analysis of the most common statistical indexes, summarised in Table 2, it can be observed that, even if graphically the data seems to have a good agreement, some indices, such as R^2, seem to suggest bad correlation. It has to be noticed that those indexes are highly affected by single extreme values. In fact, by removing a single data on 10th July 2017, that was particularly high (5 log_{10}) and thus could be considered as an outlier, R^2 increased up to 0.2. Moreover, even if the statistical model did not provide the exact value, it wouldn't have anyway allowed bathing. In fact, observing the index of correctly predicted contaminations (true positive rate), good performances can be noted.

Table 2: Evaluation of statistical model performances using simulated data from ProSe

	SC4		SC6		SC7	
R2	0.86	0.04	0.87	0.09	0.87	0.09
Mean squared error	0.06	0.38	0.06	0.4	0.06	0.45
Sample size	404	202	404	202	404	202

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

% below 95th percentile	-	93.0	-	91.0	-	92.0
% below 90th percentile	-	90.0	-	85.0	-	85.0
% in 95% prediction interval	-	94.0	-	95.0	-	95.0
Correctly predicted contaminations (True-positive rate)	-	0.99 (70/71)	-	0.92 (54/59)	-	0.91 (48/53)
Unpredicted contaminations (False-negative rate)	-	(/71)	-	(/59)	-	(/53)

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

3. DS3 Early Warning System for safe water reuse

The EWS for safe water reuse is a tool conceived within the risk-based management framework of sanitation systems. It aims at preventing bacterial and toxic contamination linked to the reuse of treated wastewater for agricultural irrigation based on:

- A comprehensive network of multi-parameter sensors at a WWTP
- New sensors for real-time and in-situ measurements (e.g., *E. coli* measurement - Digital Solution DS1)
- Machine learning and statistical correlation to assess contamination risk

The EWS that was developed for safe water reuse at Peschiera Borromeo WWTP was designed on the basis of a Risk Assessment and can be integrated with the real data acquired and processed at the plant by a dedicated SCADA system and control room, so as to potentially deliver dynamic risk management and decision support for risk minimization.

In this work, the risk assessment was performed in a semi-quantitative mode by the construction of a risk matrix to understand the main components of risk in the WWTP. Furthermore, a quantitative microbial risk assessment (QMRA) was performed for a better characterization of the microbial risk. Hence, outcomes of the risk analysis as well as regulatory limits were considered for the selection of the parameters that are needed to monitor to assure a safe wastewater reuse. Hence the monitoring network constituted by sensors and soft sensors was defined. Particularly, the EWS can be defined as an integration of sensors and soft-sensors able to generate warning and alarm (when specific thresholds for selected parameters are exceeded) to allow facility staff to react promptly and take preventive actions that will assure a safe wastewater reuse.

3.1. Planning and design: description of the solution based on WSP and SSP

The EWS developed for Peschiera Borromeo water reuse system relies on the concepts and principles typical of a Water Safety Plan (WSP) and/or a WHO Sanitation Safety Plan (SSP). WSPs provide a systematic approach towards assessing, managing and monitoring risks from catchments to drinking-water consumers. At the same time, SSP applies the same conceptual and procedural approach from sanitation waste generation to the waste's final use and/or disposal – both reclaimed water and treated sludge.

In the case of reuse/recycled waste streams in agriculture, which produce a food product, SSP goes from "toilet to the farm to table", or eventually to waste streams which are released to the environment, from "toilet to the environment".

There are, however, critical differences in the two approaches: SSP typically operates in a less defined regulatory environment, has multiple objectives, has more stakeholders and addresses risks to multiple exposure groups.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

3.1.1. Structure and methodology of Water Safety Plans and Sanitation Safety Plans

Risk management requires a description of the system defining its main components and boundaries. Usually, the system is schematised in a flow diagram, in order to get a clear but complete overview of the global picture and all the parts involved. The system is then contextualised considering the regulatory framework of standards and requirements, together with the geographic, climatic and socio-economic conditions of the local area. Once the system is defined, the next phase consists of the identification of all hazards and hazardous events, the characterisation of their likelihood, and the severity of the related consequences. Risk assessment is performed following a semi-quantitative approach, attributing weights and scores in order to quantify the considerations and the evidence emerged. Possible exposure routes and exposed groups are defined, with particular attention to vulnerable subjects. Preventive and control measures are evaluated, considering their effectiveness on risk minimisation. Treatments, technologies and behavioural measures are taken into account, considering a multi-barrier approach. Monitoring level, procedures, maintenance programs and emergency plans are implemented in the evaluation.

3.1.2. Case Studies of Water Reuse Risk Management Plan

In this chapter, some case studies about the application of Water Reuse Risk Management Plans (WRRMPs) are presented.

The first case study presented was inspired to the WSP manual provided by WHO for the drinking water system and was applied for reuse of treated wastewater in green areas at the "Universidad Nacional Autonoma de Mexico (UNAM)" (campus University City) (R3WATER, 2017). The main objective was to ensure human health protection. The health-based target was identified by the "Norma Oficial Mexicana, NOM 003 SEMARNAT 1997", which establishes the maximum contaminant limits for reuse of treated wastewater in public services. The system was schematised in eight major components, and critical control points were determined considering wastewater treatment processes, reception and storing practices, accidental or deliberated contamination events, maintenance of the distribution system, protection practices and variations due to weather conditions. As a final result of the WSSP implementation, an upgrade of the WWTP and an improvement of the physical conditions of storage tanks were performed to ensure the safety of human health. At the same time, a monitoring and maintenance program was established.

As an intermediate step of a Water Reuse Safety Plan, four case-studies within the European research project DEMOWARE (2013) (i.e., El Port de la Selva, Braunschweig, Olf Ford and Sabadell) have been investigated for health risks caused by pathogens via quantitative microbial risk assessment (QMRA). This approach calculates the probability of infection, combining the calculated concentration of pathogenic microorganisms with available dose-response relationships and end-use specific exposure scenarios. The final step consists of calculating the disability-adjusted life years (DALYs), used as an indicator of disease burden.
Particularly, two-thirds of the Braunschweig WWTP effluent (ca. 15 million m^3 per year) is used for the irrigation of 2700 ha of agricultural area. Therefore, in Braunschweig quantitative

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

microbial risk assessment (QMRA) was conducted in order to quantify the probability that the planned reuse system would be able to meet the WHO health-based target of 10^{-6} DALYs per person per year (Bedoui, 2014). The selected reference pathogens were identified in R*otavirus*, *Campylobacter jejuni*, *Cryptosporidium* and *Giardia*; while for exposure assessment three different scenarios were assessed: i) exposure of fieldworkers, ii) exposure of local/nearby residents and iii) children ingesting soil irrigated with reclaimed water. The risk assessment shows that the current measures for risk reduction are sufficient to meet the WHO benchmark for water reuse for residents (including children), and all target pathogens. However, fieldworkers have an increased work-related risk of infection, which exceeds the WHO benchmark. The WHO requirement may be satisfied by combining UV disinfection and irrigation on demand.

The Australian Guidelines for recycled water management reports a real case of a risk management plan for agriculture reuse of treated wastewater (Chow et al., 2018). 120 million of raw sewage from domestic and industrial activities enter daily the WWTP that consists of secondary treatment followed by about 20 days of lagoon storage and polishing. In order to use the effluent for irrigating commercial food crops, treatment was expanded to include coagulation, dissolved air flotation and filtration (DAFF), and disinfection. From the risk assessment, it appears that human health is mainly affected by microbial hazards, while chemical aspects of recycled water (such as chloride, sodium and nutrients concentration, and salinity) produce a risk to the environmental performance. To reduce the risks, preventive measures are implemented, and critical points are identified.

To ensure the application of best practices in water reuse, the new Portuguese policy focuses on the adoption of projects supported by a risk management framework and quality standards defined according to a fit-for-purpose approach. Rebelo and colleagues (Rebelo et al., 2018) proposed a methodology based on ISO standards 16075 that allows validating appropriate quality standards for water reuse practices and helps authorities on the decision-making process. At the same, the application of the risk assessment methodology was demonstrated in a case study, namely a vineyard irrigated with reclaimed water from an urban wastewater treatment plant (Hong et al., 2012)

3.1.3. Role of Early Warning Systems in WSPs and SSPs

According to WHO's WSPs and SSPs, EWS is a control measure used to reduce risks due to unforeseen hazardous events and keep the water chain production for potable use or reuse under control. Once validated, EWS is used for operational monitoring of treatment processes. Typically, it is a fully automated tool using on-line calibrated instruments connected to a SCADA system (that is, a control system architecture comprising computers, networked data communications and graphical user interfaces), where alarm levels are typically set to provide an early warning as well as an emergency trigger. Alarms usually recall system operators' attention to attend the plant and often start automated processes to stop supplying water into the treated water storage or directly to reuse.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

In practice, automated monitoring systems require much work due to problems with selecting reliable instruments and control systems. Most utilities will persevere in improving these systems until they are sufficiently reliable and suitable for their WSPs and SSPs. Most systems are designed to have multiple triggers to avoid supplying untreated water. For instance, systems often automatically shut systems down or switched them to standby, and usually, there are early warning alarms that would provide time for problems to be fixed before they affect customers.

In the Milan case study, the risk based EWS has been developed on the basis of the evidence, and the outputs emerged from risk assessment. According to the new EU Regulation 741/2020 on minimum requirements for water reuse, "*risk management shall comprise identifying and managing risks in a proactive way to ensure that reclaimed water is safely used and managed and that there is no risk to the environment or to human or animal health. For those purposes, a water reuse risk management plan shall be established*".

3.2. Local deployment of the solution - Milan case study

In order to plan and design a suitable Risk management for safe water reuse in Peschiera Borromeo WWTP, background information was collected and elaborated in different steps, including:

1. Definition of the operational framework for the implementation of the solution (section 3.2.1)
2. Choice of the water quality class to be produced (section 3.2.4)
3. Analysis and assessment of the WWTP efficiency and resilience (section Analysis and assessment of the WWTP efficiency and resilience
4. Risk analysis (SSP) of the production chain (section 3.2.4)
5. Quantitative risk assessment

3.2.1. Definition of the operational framework for the implementation of the solution

The definition of the operational framework of the EWS for a safe wastewater reuse at Peschiera-Borromeo WWTP was conducted considering the characterises of the WWTP, including technologies applied to treat wastewater, wastewater quality, and the agricultural practices applied in the surrounding fields.

The foundation of the rationale for the development of EWS started from the existing monitoring programs defined by the national and European Regulation on water reuse and by the selection of some specific aspects to focus on in order to stress test the reuse mechanism. In particular, the choice of on-line sensors and probes which were implemented is directly linked to the idea of stopping wastewater reuse before it may affect the health safety of exposed people.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

3.2.1.1. Treatment trains at Peschiera-Borromeo WWTP

The plant selected for the deployment of the solution is located in the municipality of Peschiera Borromeo, in Via Roma - Cascina Brusada (Figure 8). The plant serves a large urban area (Milan and neighbouring municipalities) and the Lambro River is the water body receptor of the discharged wastewater. The WWTP has a treatment capacity of 566000 P.E. and treats an average flow rate of 216.000 m^3/day.

The plant includes two treatment lines receiving wastewater from different urban areas:

• Line 1: Municipalities of Brugherio (MB), Carugate, Cassina de' Pecchi, Cernusco sul Naviglio, Cologno Monzese, Peschiera Borromeo, Pioltello, Segrate and Vimodrone.

• Line 2: Municipality of Milan and Linate district of Peschiera Borromeo

Figure 8: Location of Peschiera Borromeo WWTP

The schematic of the treatment process is reported Figure 9. Below is reported a brief description of the treatment trains of Linea 1 and Linea 2. Figure 9.

Line 1: Biological Wastewater Treatment with Activated-Sludge Process followed by tertiary treatment. The process includes the following steps:

- Coarse screening, fine screening and odour treatment;

- Grit and oil removal system;

- Primary sedimentation with two circular settlers;

- Biological oxidation in an activated sludge unit;

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

- Secondary sedimentation with 4 circular settlers;

- Tertiary treatment for nitrogen removal in BIOFOR reactors;

- Final chemical disinfection with peracetic acid.

Line 2: Biological oxidation and nitrogen removal in BIOFOR reactors

- Coarse screening, fine screening and odour treatment;

- Sand separation and oil extraction in a compact SEDIPAC combined with primary sedimentation;

- BIOFOR reactors for organic carbon and nutrient removal combined with filtration. There are 10 BIOFOR modules, including 5 dedicated to pre-denitrification and 5 voted to organic removal and nitrification.

- Final disinfection through UV lamps.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

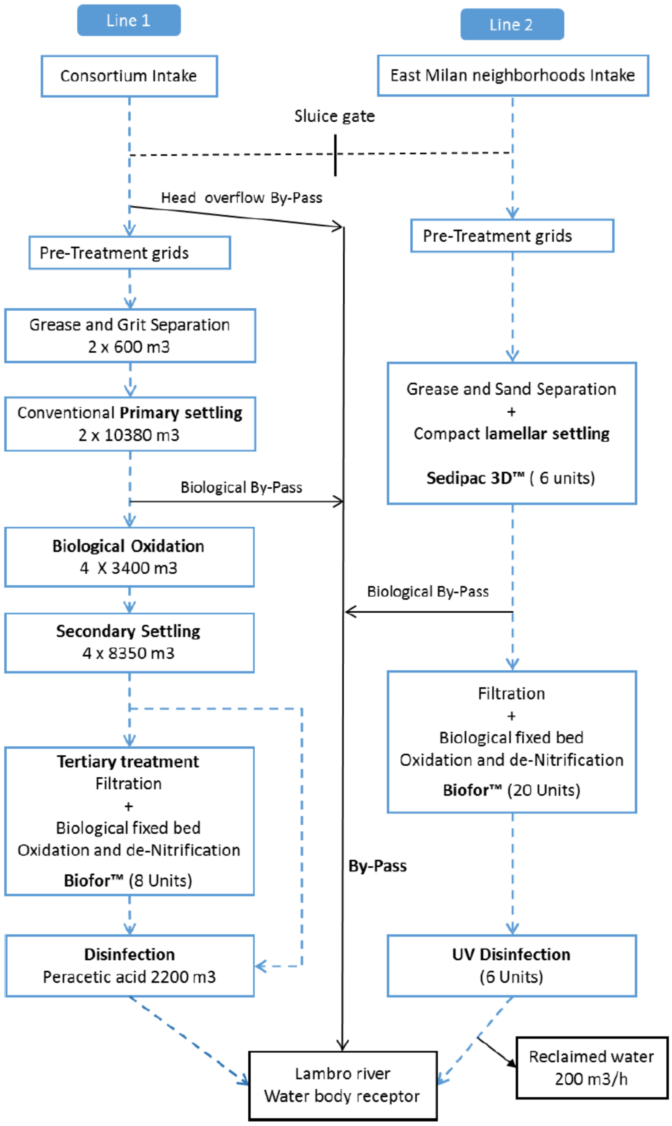


Figure 9: Detailed scheme of Line 1 and Line 2 treatment trains of Peschiera Borromeo WWTP

3.2.1.2. Monitoring Programs

At Peschiera Borromeo WWTP, monitoring of wastewater quality is performed by sensors installed at the plant and by laboratory analyses.

The digital monitoring of wastewater quality at Peschiera Borromeo WWTP is performed by a remote control and a SCADA system for the continuous acquisition of online data measured at the WWTP by sensors. Laboratory analyses are also performed periodically for influent and effluent characterization to control specific processes.

Equipment status and related alarms on electro-mechanical units are continuously monitored. Offline data about cumulative energy consumptions, chemicals supply, sludge and waste

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

production and disposal are stored in internal management systems. Maintenance operations, internal report and emergency procedures follow specific and documented protocols.

At Peschiera Borromeo WWTP, a network of sensors and probes (Figure 10) is installed at the plant. It provides on-line and real-time data to monitor the wastewater quality in the influent, during the treatment process, and in the final effluent.

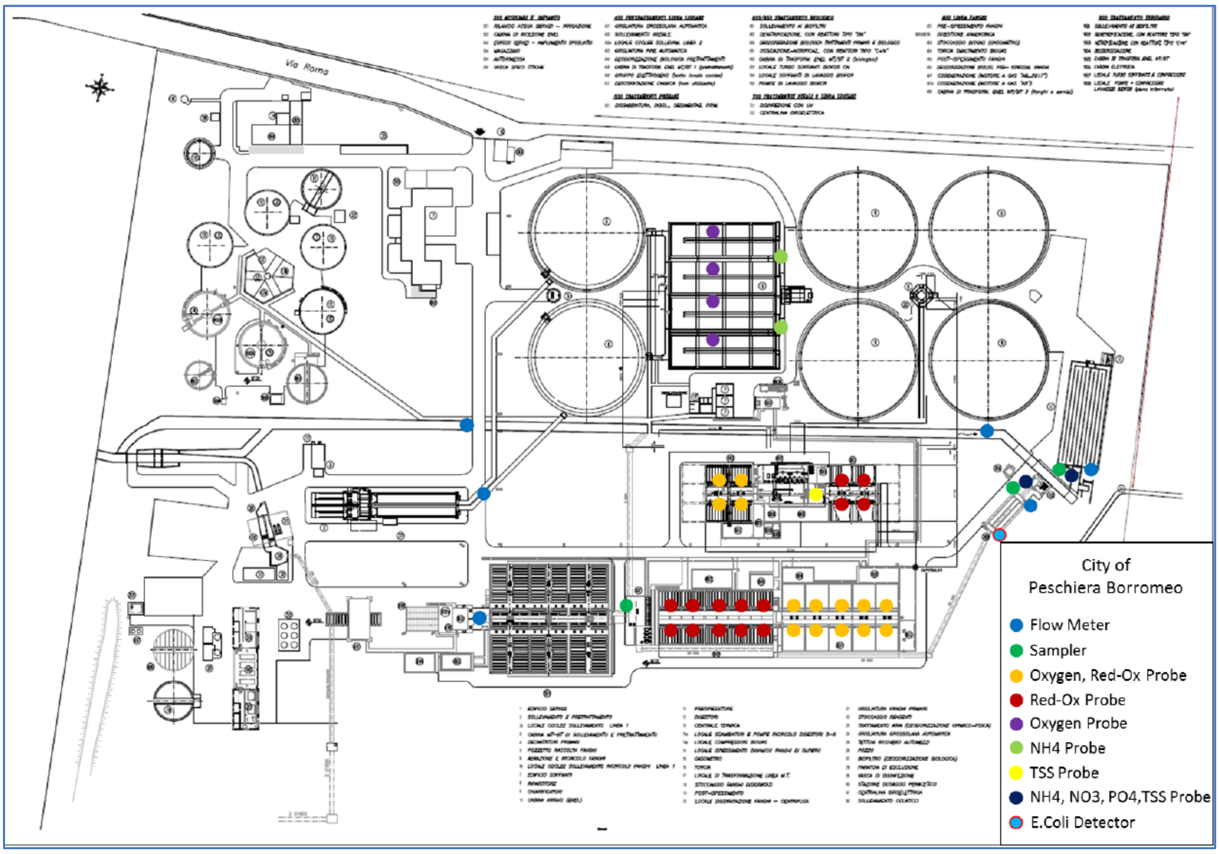


Figure 10: Layout of the treatment trains of Peschiera Borromeo WWTP

Since Line 2 was selected for the experimental activities in DWC project, below is reported a synthetic description of the main operational controls and management systems installed in this treatment line.

Screening and pumping units are equipped with level radars, and alarms in case of malfunction of the electromechanical equipment. Energy meters are installed to measure dynamically the real energy consumption. Sensors for pH, ORP, conductivity, TSS, NH_4, PO_4 are installed before the SEDIPAC unit, which is also provided with flow meters. All the equipment for sludge extraction, oil removal system and sludge conveyor are equipped with alarms. Energy meters measure the electricity consumption. On the internal back-flush, that is sent back to the SEDIPAC, chemicals are dosed for phosphorus precipitation, and the related

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

electromechanical equipment is provided with alarms. BIOFOR reactor for biologic and nutrient removal combined with filtration is divided into 10 modules, 5 dedicated to pre-denitrification and 5 voted to organic removal and nitrification. In the aerobic compartments REDOX, Temperature and Dissolved Oxygen are measured online with sensors, while the anoxic zones are provided with REDOX probes. In the internal recycle a N-NO_3 analyser is installed and the recycle flow rate is also measured. Backwashing is monitored with a flow meter, the flux is activated alternatively by temporization or by pressure signals from sensors installed on the filters surface. Energy meters are installed to monitor electricity demand. In the UV disinfection unit, sensors are installed to monitor the UV light intensity. Maintenance operations are supported by a counter system with a threshold of maximum 10000 working hours for each lamp. Specific energy meters are installed to monitor the UV disinfection unit. In the final effluent, a set of probes are installed to monitor in real time several parameters and a flow meter is installed to control the amount of treated water. Key Performance Indicators (KPIs) are automatically calculated and correlated with historical data to detect anomalies. In the final effluent, are installed sensors for the monitoring of TSS, NH_4, PO_4, TOC and UV absorbance at 254 nm in real-time.

In particular, the choice of monitoring TOC as a potential control parameter to assess the efficiency of the disinfection treatment derives from its typical use in the potable water sector, and specifically for the control of water quality in drinking water wells.

TOC analyser purchased and installed at Peschiera-Borromeo WWTP in the framework of the digital-water.city project is the model BioTector 3500 from Hach Lange. Similarly to what is commonly done in laboratory scale machines, the BioTector 3500 first analyses TIC (Total Inorganic Carbon) through sample acidification, which leads to CO_2 desorption. On a second quota of the withdrawn sample, the instrument measures TC (Total Carbon) via a chemical oxidation catalysed by hydroxy radicals, generating again a consequent flow of CO_2. The two CO_2 flows, quantified via Infrared spectroscopy, are subtracted to obtain the TOC value.

TOC along with UV absorbance at 254 nm (UV_{254}) is an indicator of organic content of the water and can provide indications to assess the performance of disinfection processes as well as risk for possible microbial regrowth (Chien et al., 2009; Masaaki K, 2014). In addition, both TOC and UV_{254} are measurements correlated to COD concentration in wastewater, and the COD to TOC ratio can give hints about characteristics of the organic matter present in the water.

Sampling and periodical laboratory monitoring

The sampling and the periodical laboratory monitoring at Peschiera Borromeo WWTP is carried out for the two wastewater lines: Line 1 and Line 2. The laboratory, which is accredited according to the UNI CEI EN ISO/IEC 17025:2018, is located inside the WWTP. Analytical results are typically available within 24h (5 days for BOD_5) and are stored by a software for data

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

management (Water LIMS). The same software allows to send emails and alerts for any parameters that exceed the regulation limits.

The monitoring program foresees the accomplishment of the analyses of several parameters, including BOD_5, COD, SST, TP, TN e NH_4, in 24 hours composite samples taken from the raw influent and the final effluent. These analyses are aimed to verify the compliance of the treated water with standard limits.

The wastewater treated in Line 2 has to be in compliance with D.M. 185/ 2003 to perform wastewater reuse. Hence, at Peschiera-Borromeo WWTP samples are analysed for the determination of parameters defined in D.M. 185/ 2003. Particularly, the outlet treated water is analysed for: COD, BOD_5, NH_4, TP, TN, TSS, metals (Al; As; Cd; Cr; Ni; Pb; Cu; Zn; Fe), pH, conductivity, chlorides, Boron, Sulphate, SAR (Sodium Adsorption Ratio), aromatic organic solvents, organic nitrogenous solvents, total surfactants. Furthermore, weekly *E. coli* concentration is measured. The list of the Chemical-physical parameters analysed at Peschiera Borromeo Laboratory is reported in Table 3.

Table 3: Chemical-physical parameters analysed at Peschiera Borromeo WWTP Laboratory

Parameter	LOD	LOQ	Accredited Method
Conductivity at 25°C		150 mS/cm	YES
BOD5		5 mg/L	YES
COD		15 mg/L	YES
TSS		5 mg/L	YES
Total nitrogen (TN)	0.03 mg/L	0.1 mg/L	YES
Ammonium nitrogen (NH_4)	0.043 mg/L	0.44 mg/L	YES
Ammonium nitrogen (N)		0.5 mg/L	YES
Nitric nitrogen (as N)	0.002 mg/L	0.049 mg/L	YES
Nitrous nitrogen (as N)		0.1 mg/L	NO
Total phosphorus (TP)	0.01 mg/L	0.313 mg/L	YES
Phosphates (PO_4)	0.002 mg/L	0.023 mg/L	YES
Aluminium (Al)	0.005 mg/L	0.112 mg/L	YES
Arsenic (As)		0.03 mg/L	NO
Barium (Ba)		0.1 mg/L	NO
Boron (B)		0.1 mg/L	NO
Cadmium (Cd)	0.0004 mg/L	0.010 mg/L	YES
Total chromium (Cr)	0.0016 mg/L	0.045 mg/L	YES
Hexavalent chromium (Cr VI)		0.003 mg/L	NO
Iron (Fe)	0.0029 mg/L	0.041 mg/L	YES
Manganese (Mn)	0.0035 mg/L	0.115 mg/L	YES
Mercury (Hg)		0.0005 mg/L	NO
Nickel (Ni)	0.0009 mg/L	0.050 mg/L	YES
Lead (Pb)	0.0023 mg/L	0.024 mg/L	YES
Copper (Cu)	0.0003 mg/L	0.010 mg/L	YES
Zinc (Zn)	0.0011 mg/L	0.055 mg/L	YES
Chloride (Cl^-)	0.018 mg/L	3.478 mg/L	YES

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Sulfate (SO_4)	0.023 mg/L	2.292 mg/L	YES
Sulphite (SO_3)		0.5 mg/L	NO
Sulfide (S)			NO
Cyanides (Cn)		0.01 mg/L	NO
Fluoride (F)		0.25 mg/L	NO
Phenols			NO
Total hydrocarbons		0.05 mg/L	NO
Animal/vegetable fats and oils		10 mg/L	NO
Non-ionic surfactants (MBAS)		0.5 mg/L	NO
Non-ionic surfactants (BIAS)		0.2 mg/L	NO
Total surfactants		0.2 mg/L	NO
Chlorinated solvents		1 mg/L	NO
Organic Aromatic Solvents		0.1 mg/L	NO
Sedimentable solids		0.1 mg/L	NO
COD after filtration 0.45 microns		15 mg/L	NO
Dry residue at 105 ºC			NO
Suspended solids 105°C		5 mg/L	NO
Total solids at 105 °C		5 mg/L	NO

3.2.1.3. Agricultural practices in the surrounding areas

In the Po valley, the agricultural irrigation systems are highly dependent on freshwater diverted mainly from rivers. Particularly, in the plain of the Lombardy region irrigation has been developed since the Middle Age, as witnessed by the extensive network of historical channels, which convey water for irrigating an area of about 550 000 ha. The irrigated area represents about 85% of the total utilised agricultural area, demonstrating the relevance of irrigation to sustain the agricultural sector of Lombardy, one of the most important in Europe, both in terms of quantity and quality of agricultural and food production. Given its geographical context, agriculture plays a key role in the economy of the region. In this area the most employed agriculture activities are fodder crops, such as maize. The most common irrigation techniques rely on border irrigation, implying a water demand of about 3000-4500 m^3/(ha*month) during the irrigation season.

Crop quality impacts both the quantity and quality of the required irrigation water because different types of crops need different amounts of water and are sensitive to different water characteristics. Furthermore, the amount of water for irrigation and the frequencies of the irrigation events vary according to the period of the year. Crop quality also implies numerous risk considerations, which depend on harvesting practices, processes accomplished to obtain the final product, and the final use of the product itself (e.g., consumed cooked or raw, consumed by humans or animals).

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

At present the most common sources of water for irrigation are rivers. Water is diverted from rivers and then distributed by gravity through a network of open, unlined canals and is supplied to farmers under rotation. The irrigation network of canals and the related water delivery systems are supervised by the consortium Est Ticino Villoresi, one of Italy's most important irrigation consortia.

Irrigation water availability has been traditionally adequate, but competition for water resources access in the area has been increasing in the last years and it is expected to further escalate, mostly due to a decreasing trend in summer water availability and to a reinforcement of the ecological flow requirements. Therefore, the identification of additional water resources, along with the improvement in irrigation efficiency, represent key issues for the next years. Alternative sources of water, such as reclaimed wastewater, are drawing attention as innovative solutions.

In the context of Digital Water City – Horizon 2020 project, a study aiming at demonstrating the potential of a smart reuse of treated wastewater was carried out. The study considers one of the two treatment lines ("Line 2"), that provides an average discharge of 1 m^3/s

The study was carried out in the agricultural season 2021 in two fields adjacent to the WWTP. Both fields are cultivated with the same agronomic practices, but one field is irrigated using the most common irrigation method (i.e., border irrigation), while in the second one an attempt to implement a "smart" irrigation practice was performed (Figure 11).

The fields are cropped with maize during the summer season, while during autumn and winter mustard is cultivated as a cover crop. In 2021 maize was sown at the beginning of April (April 8th) and harvested on September 29th. The fields have been irrigated in past years using border irrigation with a centrifugal pump that lifts the water from a canal and spreads it over the field. In one of the fields a new drip irrigation system was installed for the purpose of the DWC experimentation. The main pipe starts from the outlet of WWTP Line 2 and reaches the field boundary where a manifold is connected.

The field is divided into four different sectors where drip irrigation can be activated autonomously through four remotely controlled electro-valves. Each sector was irrigated for approximately 12 hours every two days during the agricultural season. At the beginning of the season, laterals, connected to the manifold, were installed in the crop inter-row with a spacing of 1.4 m and were partially buried; the emitters' distance and discharge are 30 cm and 1.14 l/h, respectively.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Figure 11: Overview of demo site (left), drip irrigation system installation (centre), electro-valve regulating irrigation in each sector (right)

To obtain information about soil water status, two multilevel humidity probes (Figure 12) located at two different points along the drippers line (yellow triangles in Figure 11) and a piezometric well (blue dot) with a sensor (Figure 12) to monitor the ground table depth were installed.

The same sensor configuration was used in the border irrigated field.

In addition, a weather station (Figure 12) was installed near the demo site to measure the local weather agrometeorological variables that are required to estimate the crop evapotranspiration; for security reasons, the station was installed inside the WWTP, about 500 m from the field (red star in Figure 11).

Figure 12: Sensors and devices installed: piezometer, water content probe + GSM modem (left); agro-meteorological weather station (right)

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

3.2.2. Choice of the water quality class to be produced

3.2.2.1. Legislative boundaries

Before the entry into force of European Regulation 2020/741 (hereinafter also referred as "Reuse Regulation"), the reuse of water was regulated at European level by Directives: 2000/60/EC, Art. 11 and 91/271/EEC, Art. 12, while at the Italian National level by the Decrees: 12/6/2003, n. 185 of the Ministry of the Environment, and by the Legislative Decree 3 April 2006, n. 152, art. 99. Particularly, with the decree n.185 on regulation on technical standards for the reuse of wastewater, Italy has been among the 7 EU countries with ad hoc legislation since 2003.

The main four issues to be resolved in the EU context were:

1. the lack of uniform EU legislation
2. the perceived lack of consumer protection
3. the presence of obstacles to the movement of agricultural products linked to health and environmental risks
4. the high cost of the wastewater reuse system.

The EU's response to these needs has been a detailed European framework with minimum requirements for the reuse of water for irrigation, and with the following objectives: circular economy, environmental sustainability, qualitative and quantitative protection of water.

The choice of the Regulation as a regulatory act over the Directive has had the greatest advantage in that it is an immediately enforceable act, with an entry into force on 26.06.2020 and an application from 26.06.2023. The Reuse Regulation provides a special flexibility in terms of:

1. Application is foreseen three years after entry into force
2. Substantive rules laid down in the EU Regulation
3. Administrative procedures and measures to ensure the effectiveness defined by the Member States
4. Discretion of the member states on the application of re-use; criteria laid down in the Regulation; decision to be communicated to the EC and to be reviewed every 6 years in line with the Management Plans, the climatic conditions, the pressures on the water bodies from which the water is taken and the wastewater receiving bodies and the environmental and resource costs
5. The competent authority may lay down any other additional minimum requirements, including those relating to monitoring

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

The main pillars of EU Regulation 2020/741 are the two Annexes: the first one relevant to the uses of reclaimed water and minimum requirements (Table 4) and the second one in relation to the key elements of risk management that will necessarily have to be site specific.

According to the field of application, the regulation is mandatory for irrigation in agriculture, while it is optional for industrial, civil and environmental use. While according to the regulatory framework, reclamation facility operators will have the obligation of monitoring, the competent authorities will be responsible for the compliance checks.

Table 4: EU Regulation Table 1 of Annex I on Classes of reclaimed water quality and permitted agricultural use and irrigation method

Minimum reclaimed water quality class	Crop category (*)	Irrigation method
A	All food crops consumed raw where the edible part is in direct contact with reclaimed water and root crops consumed raw	All irrigation methods
B	Food crops consumed raw where the edible part is produced above ground and is not in direct contact with reclaimed water, processed food crops and non-food crops including crops used to feed milk- or meat-producing animals	All irrigation methods
C	Food crops consumed raw where the edible part is produced above ground and is not in direct contact with reclaimed water, processed food crops and non-food crops including crops used to feed milk- or meat-producing animals	Drip irrigation (**) or other irrigation method that avoids direct contact with the edible part of the crop
D	Industrial, energy and seeded crops	All irrigation methods (***)

(*) If the same type of irrigated crop falls under multiple categories of Table 1, the requirements of the most stringent category shall apply.

(**) Drip irrigation (also called trickle irrigation) is a micro-irrigation system capable of delivering water drops or tiny streams to the plants and involves dripping water onto the soil or directly under its surface at very low rates (2–20 litres/hour) from a system of small-diameter plastic pipes fitted with outlets called emitters or drippers.

(***) In the case of irrigation methods which imitate rain, special attention should be paid to the protection of the health of workers or bystanders. For this purpose, appropriate preventive measures shall be applied.

The main purpose in risk management (Art.5 and Annex II) is to ensure the proactive and safe management of refined wastewater, without risk to human and animal health and without environmental risk. The relevant ownership is mainly on the reclamation facility operators and, depending on circumstances, on responsible parties as well as on end-users.

What is relevant is that in risk management the utilities and the reclamation facility operators have a key role in drawing up the risk management plan on which the permission given by the competent authority is based. The Regulation lays down all the elements of the risk analysis to be assessed as: waste water sources, treatment stages, technology used, delivery infrastructure, distribution and storage, intended use, place, period of use, irrigation techniques, type of crop/s, other water sources in case of mixing, treated water volumes to be supplied, responsible parties, potential hazards and possible malfunction, environments and populations at risk, environmental factors (soil type, ecology, etc.).

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

The risk management plan shall also include additional requirements (necessary to reduce risks), preventive measures, obligations also imposed by other EU legislation.

Regarding the obligations concerning the "permit" (Art. 6), the responsible parties apply for permission to the competent Authority, that releases it in short time except complexity of the appraisal. The permit shall set out the following: obligations of the reclamation facility operators and of all responsible parties, quality classes for refined water, intended use of crops, place of use, refineries, annual volume of refined water, additional requirements, period of validity of the permit, compliance point.
According to the monitoring and compliance obligations, reported in Art. 4 and Art. 7, the reclamation facility operators shall be responsible for monitoring at the point of compliance, in accordance with the minimum requirements laid down in the Regulation and any additional requirements laid down in the permit. On the other hand, the competent authority will be responsible for the verification of compliance carried out on the basis of the permit and the findings of the monitoring carried out by the utilities. In case of non-compliance, the Competent Authority will: (i) order the Recovery Measures Manager, (ii) inform end-users, (iii) suspend the supply of refined water until compliance is restored (in case of significant risk).
Another important aspect of the Reuse Regulation refers to the Information obligations set out in articles 9, 10 and 11. These articles provide for public information on: awareness-raising, water quantity and quality, water volume, permits, results of compliance checks, contact points for cooperation between Member States and information to the European Commission, the European Environment Agency and the European Centre for Disease Prevention and Control.

3.2.2.2. Preliminary studies to determine chemical and microbiological characteristics of the treated wastewater

According to the SSP approach and to the new EU Regulation 741/2020, CAP has started for some years an investigation on the possible presence of emerging contaminants and emerging pathogens in the wastewater treated at Peschiera Borromeo WWTP. It is a precautionary measure to deep the knowledge for a safety use of the wastewater in agriculture.

Emerging Contaminants in raw and treated wastewater

In order to undertake an investigation on the presence of contaminants of emerging concern (CECs) in raw and treated wastewater from both Lines 1 and 2 of the plant of Peschiera Borromeo, the analytical laboratory of the Instituto Mario Negri carried out two sampling campaigns on three points for each line (i.e., at inlet before grit removal, at inlet before disinfection and at outlet after disinfection) as shown in Figure 13.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Figure 13: Sampling points of 2019-2021 campaigns from Line 1 (disinfection process by Peracetic Acid) and Line 2 (disinfection process by UV irradiation) for determination of emerging contaminants and emerging pathogens in raw and treated wastewater

Originally, two sampling campaigns were scheduled, one in Winter 2019 and one in Spring 2020. The sampling campaign scheduled in Winter 2019 was regularly accomplished. On the contrary, the second sampling campaign was rescheduled for the Summer 2020 due to COVID-19 pandemic. Particularly, it was accomplished the quantitative determination of 18 different molecules among pharmaceuticals (antibiotics, anti-inflammatory, anti-hypertensive, cardiovascular, CNS drug, diuretics, gastrointestinal, lipid regulators) and personal care products (PCPs). The removal efficiencies of these micropollutants in both the treatment lines (i.e., Line 1 and Line 2) are reported in Figure 14 and in Figure 15 for the first and the second sampling campaign, respectively.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

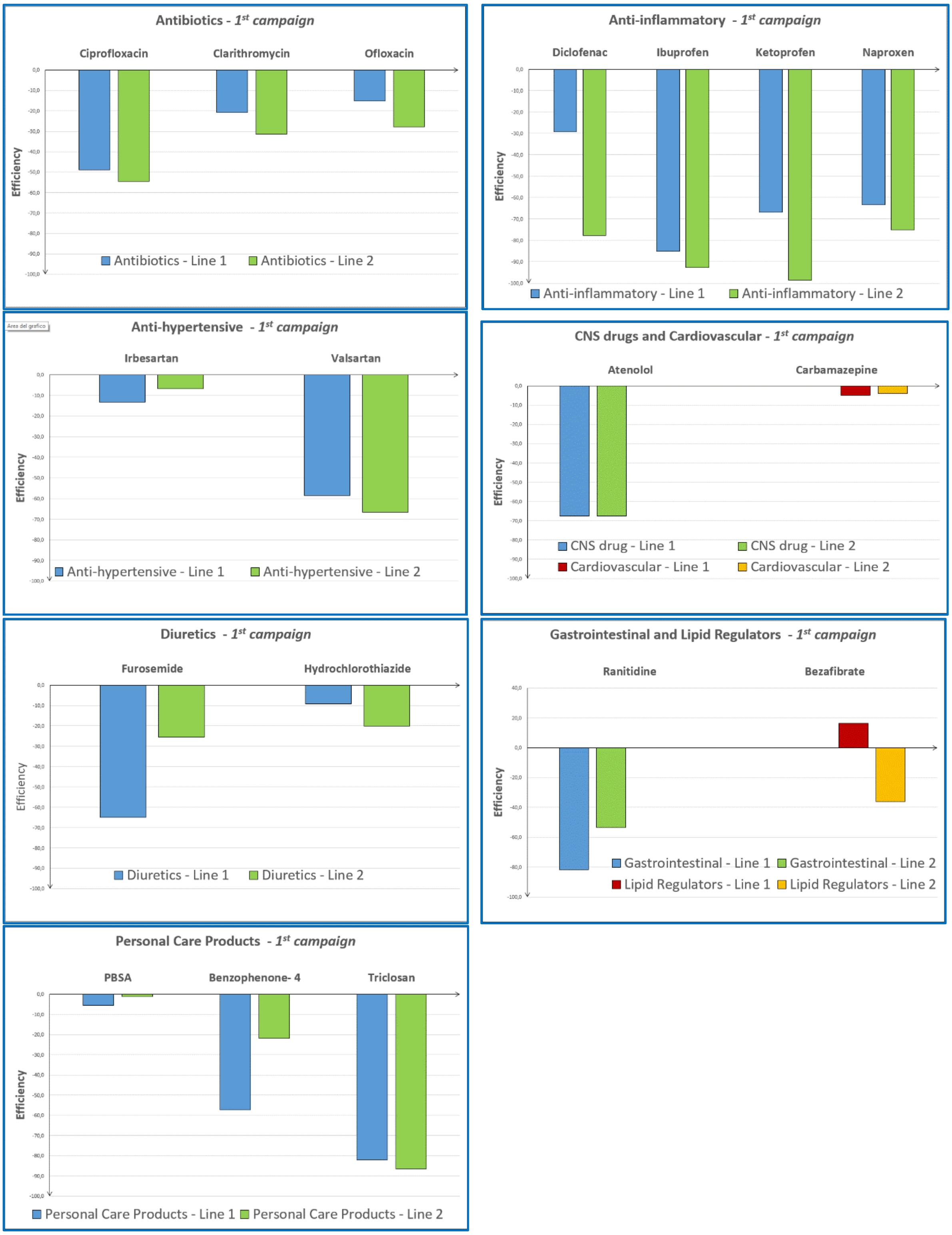


Figure 14: Comparison of drug removal efficiency of Line 1 (blue/red) respect to Line 2 (green/yellow) during the 1st sampling campaign

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

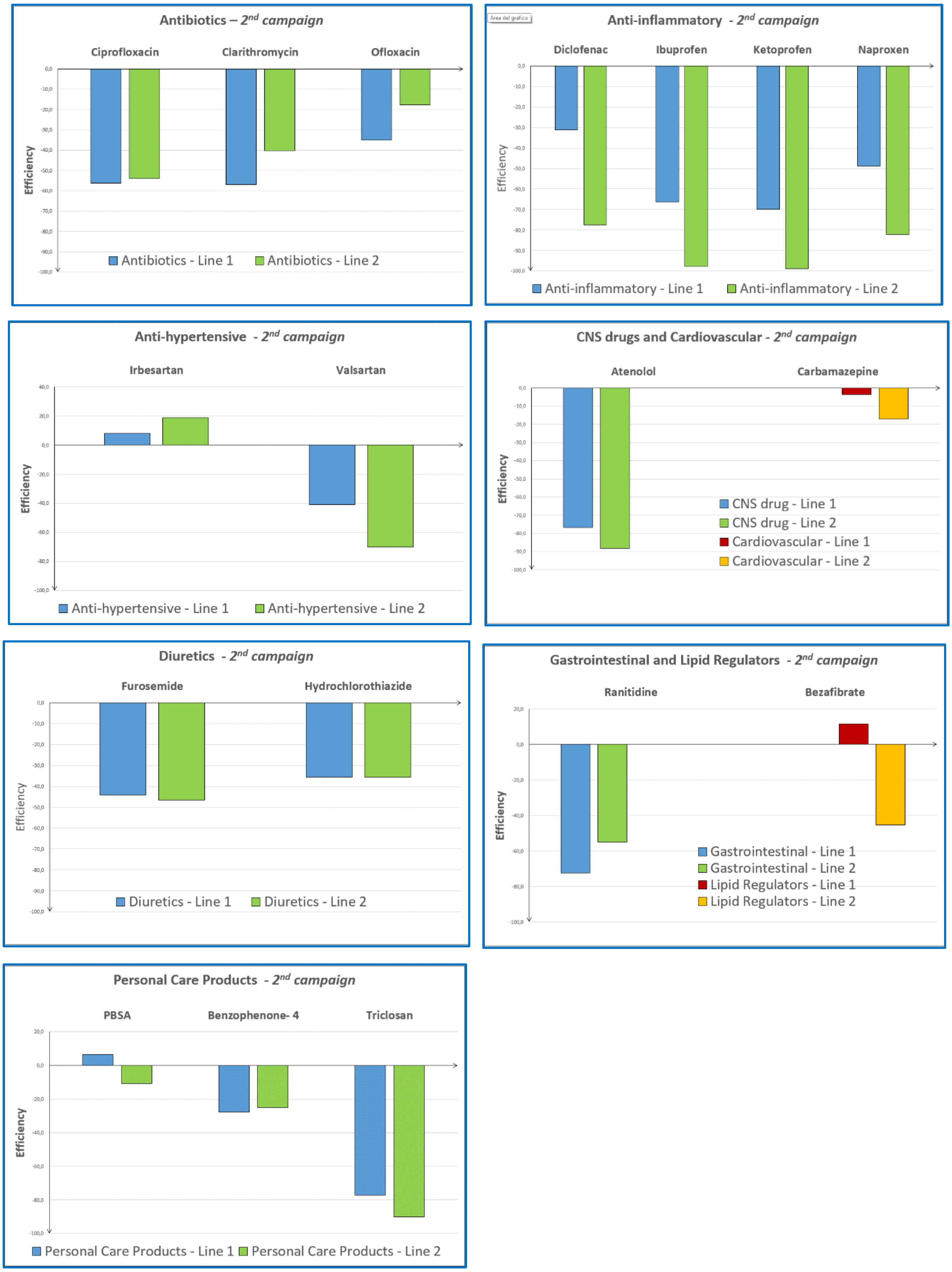


Figure 15: Comparison of drug removal efficiency of Line 1 (blue/red) respect to Line 2 (green/yellow) during the 2nd sampling campaign

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

By the analysis of Figure 14 and in Figure 15, it seems that there is not a substantial difference in the removal of the monitored micro-pollutants by the treatments processes implemented in Line 1 and Line 2 at Peschiera-Borromeo WWTP. . Among the selected substances, only ciprofloxacin and clarithromycin are part of the European Surface water Watch List. To date, there are not regulatory limits for all the analysed compounds.
In addition, according to ISO 16075-1:2020, there is no evidence today of adverse effects of these contaminants (i.e., pharmaceuticals and personal care product residuals) on human health or environment via irrigation with treated wastewater or via the consumption of crops irrigated with reclaimed wastewater.

Pathogens in raw and treated wastewater

In order to undertake an investigation on pathogens in raw and treated wastewater from both lines 1 and 2, the ISS microbiology laboratory carried out two sampling campaigns in 2021: the first was held in May 2021, the second in November 2021. As for the two sampling campaigns described in the previous paragraph, in this investigation three sampling points were selected for each line: at inlet before grit removal, before disinfection and at outlet after disinfection (Figure 13).

The selected reference microorganisms, including those pathogens belonging to the group of enteric bacteria, viruses and parasitic protozoa, were the following:

- *Campylobacter* and *Salmonella* for enteric bacteria (Bonetta et al., 2022; Santiago et al., 2018).
- *Giardia* and *Cryptosporidium* for parasitic protozoa (Bonadonna et al., 2002; Briancesco & Bonadonna, 2005a; Vernile et al., 2009).
- *Norovirus* (Katayama & Vinjé, 2018), *Adenovirus* (Allard & Vantarakis, 2018) and *Enterovirus* (Betancourt & Shulman, 2017) for enteric viruses;
- *Escherichia coli* and coliform bacteria for microbial indicators (J. B. De Souza et al., 2015; Hassaballah et al., 2019)
- *Somatic coliphages* for viral indicators (Lin & Ganesh, 2013; Truchado et al., 2021).

Samples were analysed using the following methods:

Campylobacter* and *Salmonella: (wastewater: 100 mL; pre- and post-treated/disinfected water: 1 L) Membrane filtration technique, enrichment culture, culture and isolation on selective solid media, confirmation by biochemical tests, according to ISO 17995:2019 and ISO 19250:2010 standards, respectively**.**

***Escherichia coli* and coliform bacteria:** (appropriate sample dilutions until 100 mL): Miniaturized MPN Method according to ISO 9308-2:2012 standard.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Giardia and Cryptosporidium: sampling and analysis were conducted according to ISO 15553:2006 standard. The sampling was conducted through *in loco* filtration of variable water volumes (wastewater: 2 L; pre-and post-treated/disinfected water: 50 L), using a peristaltic pump.
Filtration was performed using compressed foam filter modules (Filta-Max xpress Filter Modules, Idexx) that were eluted at lab by Filta-Max xpress pressure Elution Station (Idexx), showed in Figure 16. After concentration by centrifugation, pellets were clarified by immunoseparation with Dynabeads anti-*Giardia* and anti-*Cryptosporidium* (Dynal, Thermo Fisher Scientific). Finally, samples were stained with fluorescent-labeled antibodies (Merifluor *Cryptosporidium/Giardia* kit, Meridian Bioscience) and cysts and oocysts enumerated with an epifluorescent microscope (Zeiss), taking into consideration morphology, size and colour of the particles with respect to a positive standard.

Somatic coliphages were detected and enumerated by plaque assay, according to ISO 10705-2:2000 standard. The neutralized eluates from electropositive filters used for virus concentration (12–50 L, depending on the type of sample) were spiked on a specific growth medium, and lysis plaques were counted on a double-agar layer with bacterial host strain, after an overnight incubation.
Enteric viruses: Human Norovirus (genogroups GI and GII), Adenovirus, and Enterovirus were selected among the wide number of enteric viruses commonly found in aquatic environments, since they are excreted in large numbers by infected individuals and recognized as important waterborne pathogens. Molecular methods were used to detect and/or quantify enteric viruses: classical nested/eminested PCR followed by Sanger sequencing for all the studied viruses, and real time RT-qPCR limited to samples found positive for Norovirus GI and GII by nested PCR. For viral sampling and concentration, electropositive filter cartridges in specific housings were connected to a pumping equipment composed by a pump, a volume counter, a flow counter and a manometer, according to the EPA Method 1615. Raw/treated sewage was collected manually with a sink and poured in a cask from which the liquid was pushed through the filtering device. After filtration, the devices were kept at 4°C in thermal bags till lab processing. Within 24 hours the cartridges were eluted adding to the housings 400 mL of a 3% Beef Extract solution pH 9.5 0.05 N glycine. Filters were then shaken on an orbital shaker for 20’ to facilitate viral particles detachment. After neutralization, the eluate was centrifuged at 4000 g for 20’ and pellet re-suspended with 5 mL 1N Na_2PO_4 pH 9. The suspension was then neutralized and stored at -20°C. Genome extraction was performed with a protocol based on the magnetic beads technology, an enhanced magnetic silica version of the BOOM technology, starting from 5 mL suspension to get a final volume of 100 µL of genome extract, stored at -80°C for future use. Nested/eminested PCR and real time qPCR were performed with protocols available in literature (Jothikumar et al., 2005; Lu & Erdman, 2006; Pina et al., 1998; UNI EN ISO 15216-2, 2019).

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Figure 16: Filta-Max xpress system for parasitic protozoa concentration and elution

Figure 17: In situ primary filtration with collecting units for viruses (left) and protozoa (right)

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

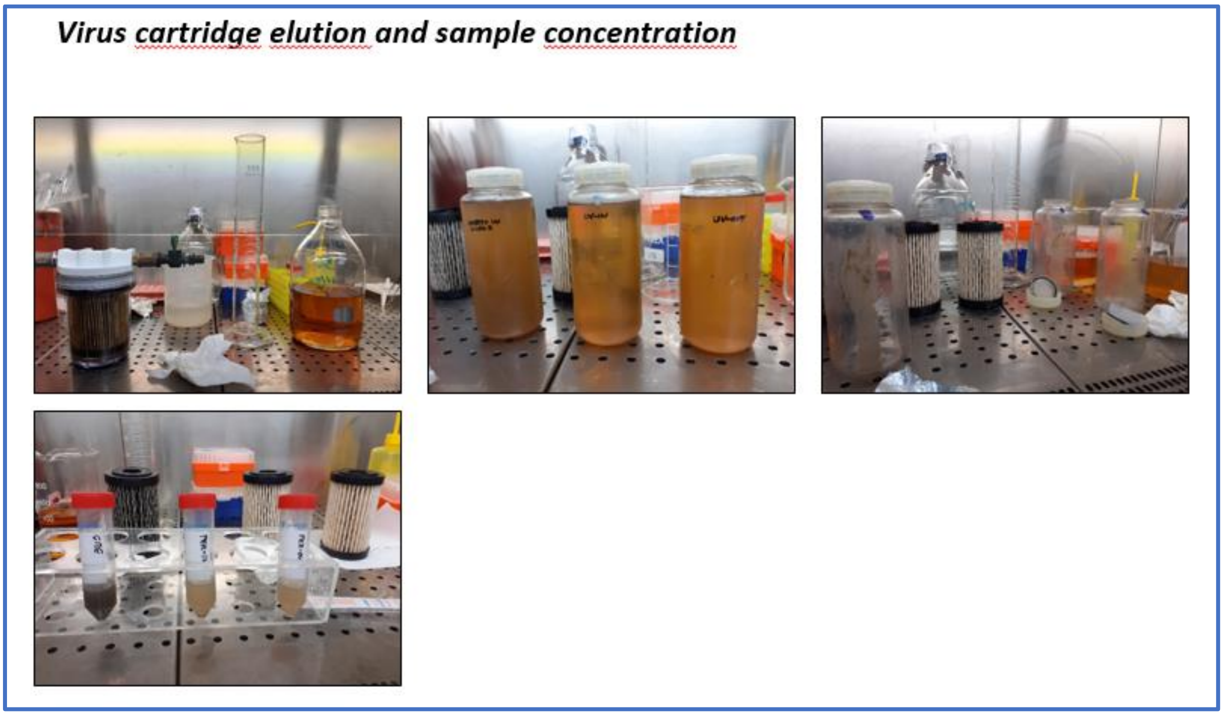


Figure 18: Elution and concentration of filtrates from samples collected in Peschiera Borromeo Wastewater Treatment Plant

The results from this microbiological screening were used as preliminary data for the validation of the monitoring strategy of the reclaimed water for agricultural irrigation from Line 1 and Line 2. Particularly, results were compared with EU Regulation 741/2020 performance targets for the treatment chain (log_{10} reduction) (Annex I, Table 4, reported in the text as Table 5). The references were the most stringent requirements for reclaimed water class (quality Class A, Annex I, Table 4).

Table 5: EU Regulation of Annex I on Validation monitoring of reclaimed water for agricultural irrigation

Reclaimed water quality class	Indicator microorganisms (*)	Performance targets for the treatment chain (log_{10} reduction)
A	*E. coli*	≥ 5,0
	Total coliphages/F-specific coliphages/somatic coliphages/coliphages (**)	≥ 6,0
	Clostridium perfringens spores/spore-forming sulfate-reducing bacteria (***)	≥ 4,0 (in case of *Clostridium perfringens* spores) ≥ 5,0 (in case of spore-forming sulfate-reducing bacteria)

(*) The reference pathogens *Campylobacter*, Rotavirus and *Cryptosporidium* may also be used for validation monitoring purposes instead of the proposed indicator microorganisms. The following log_{10} reduction performance targets shall then apply: *Campylobacter* (≥ 5,0), Rotavirus (≥ 6,0) and *Cryptosporidium* (≥ 5,0).

(**) Total coliphages is selected as the most appropriate viral indicator. However, if analysis of total coliphages is not feasible, at least one of them (F-specific or somatic coliphages) shall be analysed.

(***) *Clostridium perfringens* spores is selected as the most appropriate protozoa indicator. However, spore-forming sulfate-reducing bacteria are an alternative if the concentration of *Clostridium perfringens* spores does not make it possible to validate the requested log_{10} removal.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

From the following results of microbiological and virological analyses it was possible to obtain some preliminary considerations.

Table 6: Results of microbiological analysis (1st sampling - May 2021)

Sample	Coliform Bacteria	*E. coli*	*Coliphages/somatic*	*Salmonella spp.*	*Campylobacter spp.*
Line 1	MPN/100 mL	MPN/100 mL	UFP/ filtered volume	MPN/analyzed volume	MPN/analyzed volume
Raw Wastewater	3,6 x 10^6	2,4 x 10^5	4,2 x 10^4/12 L	11,6/100 mL	analysis not performed
Pre-disinfection	3,2 x 10^3	4 x 10^2	2,8 x 10^3/50 L	2,7 x 10^1/ 1 L	analysis not performed
Post-disinfection	2,5 x 10^3	5,3 x 10^2	4 x 10^2/50 L	0 /1 L	analysis not performed
Line 2					
Raw Wastewater	*	*	8 x 10^4/50 L	*	*
Pre-disinfection	2,4 x 10^3	1,3 x 10^2	4 x 10^2/50 L	4,5/1 L	analysis not performed
Post-disinfection	7,1 x 10^1	5	0/50 L	0/1 L	analysis not performed

* Sample not taken

Table 7: Results of microbiological analysis (2nd sampling - November 2021)

Sample	Coliform Bacteria	*E. coli*	Coliphages/somatic	*Salmonella spp.*	*Campylobacter spp.*
Line 1	MPN/100 mL	MPN/100 mL	UFP/ filtered volume	MPN/analyzed volume	MPN/analyzed volume
Raw Wastewater	2 x 10^7	2 x 10^6	2,4 x 10^5/25L	43,8/ 100 mL	<1/500mL
Pre-disinfection	7,7 x 10^3	1,2 x 10^3	3,2 x 10^3/ 50 L	1,9 x 10^1/ 1 L	<1/ 1L
Post-disinfection	1,7 x 10^3	2,1 x 10^2	1,6 x 10^3/ 53 L	0/ 1 L	<1/ 1L
Line 2					
Raw Wastewater	2 x 10^7	5,5 x 10^6	3,6 x 10^5/ 25 L	4/ 100 mL	<1/ 1L
Pre-disinfection	3,2 x 10^3	6 x 10^2	6,1 x 10^4/ 56 L	0,05 / 1 L	<1/ 1L
Post-disinfection	5,5 x 10^2	6,1 x 10^1	4,4 x 10^2/ 50L	0/ 1 L	<1/ 1L

Table 8: Results of parasitic protozoa analysis (1st sampling - May 2021)

Sample	*Giardia* n. cisti/100 L	*Cryptosporidium* n. oocisti/100 L
Linea 1		
Raw Wastewater	6 x 10^4	6,5 x 10^2
Pre-disinfection	1,7 x 10^3	4 x 10^1
Post-disinfection	1,2 x 10^2	4 x 10^1
Linea 2		
Raw Wastewater	6,5 x 10^4	2,3 x 10^2
Pre-disinfection	1,2 x 10^3	3 x 10^1
Post-disinfection	6 x 10^1	1 x 10^1

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Table 9: Results of parasitic protozoa analysis (2nd sampling - November 2021)

Sample	*Giardia* n. cisti/100 L	*Cryptosporidium* n. oocisti/100 L
Linea 1		
Raw Wastewater	1×10^4	$1{,}2 \times 10^2$
Pre-disinfection	1×10^2	2
Post-disinfection	$1{,}3 \times 10^2$	0
Linea 2		
Raw Wastewater	$1{,}7 \times 10^4$	$5{,}9 \times 10^2$
Pre-disinfection	3×10^1	4×10^1
Post-disinfection	5×10^1	7

The performance target for *E. coli* (≥ 5,0 $\log_{10}$ reduction) was achieved only at UV disinfection unit (line 2). *Campylobacter* was never found in samples from both treatment lines. Finally, oocysts were always found upstream and downstream of both treatment lines (except one case in line 1, where was observed a total removal of *Cryptosporidium*) but the removal achieved was lower than performance targets indicated by the EU regulation on water reuse.

In one case somatic coliphages were totally removed at the line 2 (UV treatment). This condition complies with EU Regulation which states that *if a biological indicator is not present in sufficient quantity in raw wastewater to achieve the $\log_{10}$ reduction, the absence of such biological indicator in reclaimed water shall mean that the validation requirements are complied with*. On the contrary, this indicator organism was always found upstream and downstream at the treatment line 1 (i.e., peracetic treatment) with removal efficiency not complying to the required performance targets for somatic coliphages (≥ 6 $\log_{10}$) in class A wastewater.

Table 10: Results of qualitative viral analysis by conventional PCR (presence/absence) in wastewater collected at Peschiera Borromeo WWTP during the two collecting campaigns (positive samples were confirmed and characterized by Sanger sequencing). EV=Enterovirus; HAdv=Human adenovirus; HNoV=Human Norovirus (1st sampling - Spring 2021)

May 10 - 11 2021 Sampling	*Conventional RT-PCR*			
Sample	***Enterovirus***	***Adenovirus***	***Norovirus GI***	***Norovirus GII***
Line 1				
Raw Wastewater	+ non polio EV	-	+ HNoV GI.1	HNoV GII.3
Pre-disinfection	-	-	-	-
Post-disinfection	-	-	-	-
Line 2				
Raw Wastewater	+ non polio EV	+ HAdV sp. D	+ HNoV GI.3	HNoV GII.3
Pre-disinfection	+ non polio EV	-	+ HNoV GI.1	-
Post-disinfection	-	-	-	-

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Table 11: Results of qualitative viral analysis by conventional PCR of wastewater collected at Peschiera Borromeo WWTP during the two collecting campaigns (positive samples were confirmed and characterized by Sanger sequencing). EV =Human Enterovirus; HAdV=Human Adenovirus; HNoV=Human Norovirus (2nd sampling - Fall 2021)

November 10 - 11 2021 Sampling	*Conventional RT-PCR*			
Sample	***Enterovirus***	***Adenovirus***	***Norovirus GI***	***Norovirus GII***
Line 1				
Raw Wastewater	-	-	-	HNoV GII.3
Pre-disinfection	-	-	-	HNoV GII.3
Post-disinfection	-	-	-	-
Line 2				
Raw Wastewater	-	-	-	HNoV GII.3
Pre-disinfection	-	-	-	HNoV GII.3
Post-disinfection	-	-	+ HNoV GI.1	HNoV GII.3

Table 12: Results of viral quantification by RT-qPCR of wastewater collected at Peschiera Borromeo WWTP for Norovirus GI and GII positive samples (1st sampling - Spring 2021)

May 10 - 11 2021 Sampling	*RT-qPCR*	
Sample	***Norovirus GI (gc/L)***	***Norovirus GII (gc/L)***
Line 1		
Raw Wastewater	3.93E+01	5.20E+02
Pre-disinfection	-	-
Post-disinfection	-	-
Line 2		
Raw Wastewater	2.28E+01	3.34E+03
Pre-disinfection	6.4E+00	-
Post-disinfection	-	-

gc=genome copies

Table 13: Results of viral quantification by RT-qPCR of wastewater collected at Peschiera Borromeo WWTP for Norovirus GI and GII positive samples (2nd sampling - Fall 2021)

November 10 - 11 2021 Sampling	*RT-qPCR*	
Sample	***Norovirus GI***	***Norovirus GII***
Line 1		
Raw Wastewater	-	3.26E+04
Pre-disinfection	-	1.53E+02
Post-disinfection	-	-
Line 2		
Raw Wastewater	-	3.26E+05
Pre-disinfection	-	2.77E+04
Post-disinfection	1.94E+01	1.71E+01

gc=genome copies

Results on qualitative viral analysis are reported in Table 10 and

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Table 11. Genetic characterization confirmed the presence of genomic traces of viruses in untreated wastewater with differences between the two sampling campaigns. Spring samples revealed the occurrence of all the investigated pathogens, with a prevalence of Norovirus (GI, genotype 1 and 3; GII, genotype 3) and Enterovirus (non-polio enterovirus). Only one sample (inlet line 2) was found positive for Adenovirus species D. No serotype has been reported in the closest prototype in genebank. No positives were observed in pre- and post-disinfection sampling points of line 1, while the pre-disinfection point of line 2 was positive for non-polio Enterovirus and Norovirus GI.1. Wastewater samples collected in fall gave positive results only for Norovirus with a higher prevalence of genogroup GII.3, which was found in all samples with exception of the post-UV disinfection point where Norovirus GI.1 was detected.

Quantitative viral analysis results are reported in Table 12 and Table 13. They were performed only on Norovirus-positive samples found by conventional PCR and revealed up to 1-2 log reduction of genome copies along the two treatment lines (proceeding from inlet to final disinfection). Despite spring outlet (post-disinfection) samples appeared to be free of viral contamination. Samples collected in fall showed still a slight positivity for Norovirus GI. Our findings are consistent with previous studies that detected norovirus in treated wastewaters (Katayama & Vinjé, 2018). It is known that viruses are ubiquitous and persistent in raw and treated wastewater, and that WWTP processes are not completely effective in the reduction of viral genomes concentrations for most viruses in wastewater (Corpuz et al., 2020). However, it is important to note that only molecular methods have been used in this study. These methods are able of detecting only viral genomes and cannot provide information on infectivity. Thus, positive results obtained in the present study do not necessarily indicate an actual threat to human health. Wastewater is indeed a very hostile environment that represents a challenge for viruses to preserve their integrity and infectivity. Results of this study have been performed on a small number of samples, and consequently should be considered as preliminary results. Indeed, the occurrence of genetic traces of Norovirus GI in post-disinfected water, at very low level, do not exclude possible residual genomic contamination after disinfection. On the other side, the lack of genomic traces in untreated and pre-disinfected wastewater could be explained by the presence of inhibitors that could affect the PCR reaction efficiency especially when the viral load is supposed to be low. Further studies are therefore needed to investigate fluctuations in viral concentration in raw and treated wastewater and the removal efficiency after treatments.

3.2.2.3. The choice of water quality class at Peschiera Borromeo WWTP

As explained in paragraph *Legislative boundaries*, European Regulation 741/2020 on water reuse provides the minimum requirements for water reuse in agriculture according to different reclaimed water quality classes, agricultural uses and irrigation methods. In the district where is located the Peschiera Borromeo WWTP, typical cultivated crops include fodder crops and crops for biomass production, while the type of irrigation is almost exclusively

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

surface irrigation. According to EU Regulation Table 1 of Annex I on *Classes of reclaimed water quality and permitted agricultural use and irrigation method* (Table 4), the quality class to select for reuse would be “Class C”. However, as precautionary measure to reduce the health risk and due to the preliminary results discussed, CAP holding S.p.A. decided to produce wastewater able to guarantee the compliance with standard limits for the water quality class B.

The minimum requirements for **class B** water quality are set out in EU Regulation Table 2 of Annex I, point (a), (Table 14).

Table 14: Reclaimed water quality requirements for agricultural irrigation according to EU Regulation 741/2020

Reclaimed water quality class	Indicative technology target	Quality requirements				
		E. coli (number/100 ml)	BOD_5 (mg/l)	TSS (mg/l)	Turbidity (NTU)	Other
A	Secondary treatment, filtration, and disinfection	≤ 10	≤ 10	≤ 10	≤ 5	*Legionella* spp.: < 1 000 cfu/l where there is a risk of aerosolisation Intestinal nematodes (helminth eggs): ≤ 1 egg/l for irrigation of pastures or forage
B	Secondary treatment, and disinfection	≤ 100	In accordance with Directive 91/271/EEC (Annex I, Table 1)	In accordance with Directive 91/271/EEC (Annex I, Table 1)	-	
C	Secondary treatment, and disinfection	≤ 1 000			-	
D	Secondary treatment, and disinfection	≤ 10 000			-	

Therefore, **class B** reclaimed water shall be considered to be in compliance with the limits set out in Table 14, when the measurements of the indicated parameters meet the following criteria:

— the indicated values for *E. coli*, *Legionella spp.* and intestinal nematodes are met in 90 % or more of the samples;

— none of the values of the samples exceed the maximum deviation limit of 1 log unit from the indicated value for *E. coli* and *Legionella spp.* and 100 % of the indicated value for intestinal nematodes;

In Table 15 are reported the requirements of the Directive 91/271/EEC, Annex I, Table1, which are also required by the EU regulation on water reuse (Table 14). In Table 16 is reported the frequency for the accomplishment of the required measurements, which requires at least 24 determinations for BOD5 and TSS parameters in the case of Peschiera-Borromeo WWTP.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Table 15: Directive 91/271/EEC (Annex I, Table1: Requirements for discharges from urban wastewater treatment plants subject to Articles 4 and 5 of the Directive. The values for concentration or the percentage of reduction shall apply.)

Parameters	Concentration	Minimum percentage of reduction (1)	Reference method of measurement
Biochemical oxygen demand (BOD5 at 20 °C) without nitrification (2)	25 mg/l O_2	70-90 40 under Article 4 (2)	Homogenized, unfiltered, undecanted sample. Determination of dissolved oxygen before and after five-day incubation at 20 °C ± 1 °C, in complete darkness. Addition of a nitrification inhibitor
Chemical oxygen demand (COD)	125 mg/l O_2	75	Homogenized, unfiltered, undecanted sample Potassium dichromate
Total suspended solids	35 mg/l (3) 35 under Article 4 (2) (more than 10 000 p.e.) 60 under Article 4 (2) (2 000-10 000 p.e.)	90 (3) 90 under Article 4 (2) (more than 10 000 p.e.) 70 under Article 4 (2) (2 000-10 000 p.e.)	— Filtering of a representative sample through a 0,45 µm filter membrane. Drying at 105 °C and weighing — Centrifuging of a representative sample (for at least five mins with mean acceleration of 2 800 to 3 200 g), drying at 105 °C and weighing

(1) Reduction in relation to the load of the influent.
(2) The parameter can be replaced by another parameter: total organic carbon (TOC) or total oxygen demand (TOD) if a relationship can be established between BOD5 and the substitute parameter.
(3) This requirement is optional.

Table 16: Minimum frequencies for routine monitoring of reclaimed water for agricultural irrigation

Reclaimed water quality class	Minimum monitoring frequencies					
	E. coli	BOD_5	TSS	Turbidity	*Legionella* spp. (when applicable)	Intestinal nematodes (when applicable)
A	Once a week	Once a week	Once a week	Continuous	Twice a month	Twice a month or as determined by the reclamation facility operator according to the number of eggs in waste water entering the reclamation facility
B	Once a week	In accordance with Directive 91/271/EEC (Annex I, Section D)	In accordance with Directive 91/271/EEC (Annex I, Section D)	-		
C	Twice a month			-		
D	Twice a month			-		

In addition to the above requirements for **class B** reclaimed water, it must be said that risk assessment for the production chain of water reuse in this project was implemented according to part C “Preventive measures” of ANNEX II of EU Regulation. In addition, some references were taken from the ISO 16075-1:2020, and supplemented by references taken from Quaderno ARSIA 5-2004 (Malorgio, 2004). These guidelines suggest parameters of wastewater quality to take into account to implement agricultural reuse, including:

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

- agronomic parameters: nutrients (nitrogen, phosphorus and potassium), salinity factors (total salt content, chloride, boron, and sodium concentration) and heavy metals' concentration
- pathogen presence

Indeed, the above-mentioned parameters can have possible impacts on crops, soil, and public health.

3.2.3. Analysis and assessment of the WWTP efficiency and resilience

3.2.3.1. Statistical analysis of collected data from laboratory monitoring program

In order to assess the efficiency and resilience of wastewater treatment Line 1 and Line 2 at Peschiera Borromeo WWTP to provide wastewater in compliance with limits for water reuse, a series of statistical analyses were carried out on the basis of laboratory data. Lab data were collected between 2018 and 2021, before and after the revamping phase of the plant. Actually, an improvement of the quality of the final effluent was observed after the conclusion of the upgrading of the plant.

To assess the continuous compliance with the EU Regulation of the treated wastewater, an evaluation of the cumulative frequencies of quality requirements achieved in each water quality class for *E. coli*, BOD_5, and TSS, for both Lines 1 and 2 of Peschiera Borromeo WWTP was conducted. In particular, in the case of the quality class B, it resulted that the quality requirement for *E. coli* was achieved for almost 73% of the analysed samples at Line 1 and for 70% of the collected samples from Line 2. These results are referred to collections accomplished from 2018 to 2021 (Figure 19). For the samples collected from 2020 to 2021, and thus, after the revamping of the plant, compliance for E. Coli concentration in class B was observed for almost 77% of samples collected in Line 1 and 87% of samples collected from Line 2 (Figure 20). On the contrary, the quality requirements for BOD_5 and TSS in class B were achieved for almost 100% of the samples collected in both Line 1 and 2 regardless of the observation period (Figure 21, Figure 22, Figure 23 and Figure 24).

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

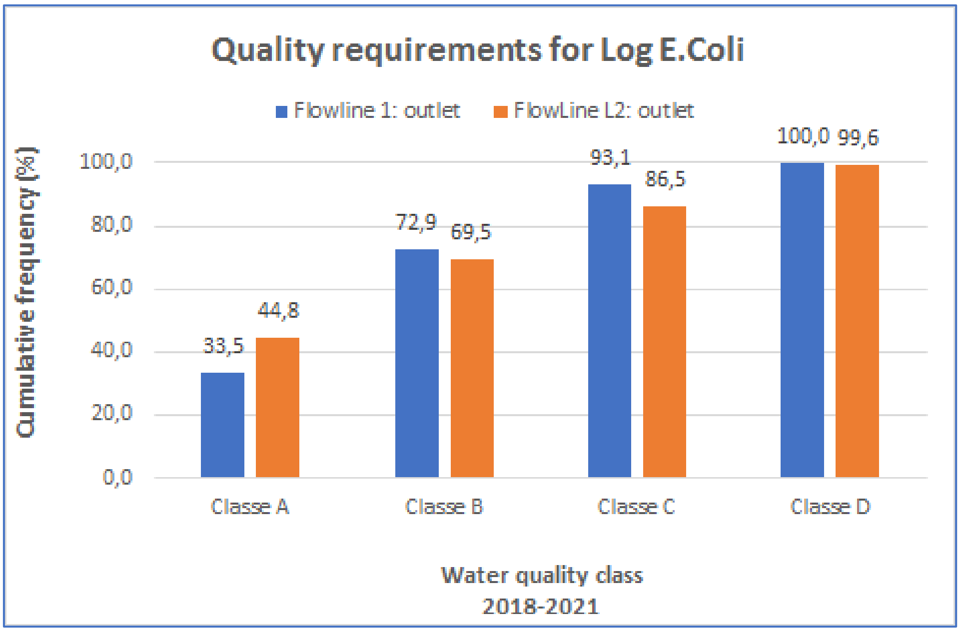


Figure 19: Cumulative frequencies from 2018 to 2021 of the quality requirements for *E. coli* achieved in each water quality class, in compliance with the recent EU Regulation on water reuse, both for Lines 1 and 2 of Peschiera Borromeo WWTP

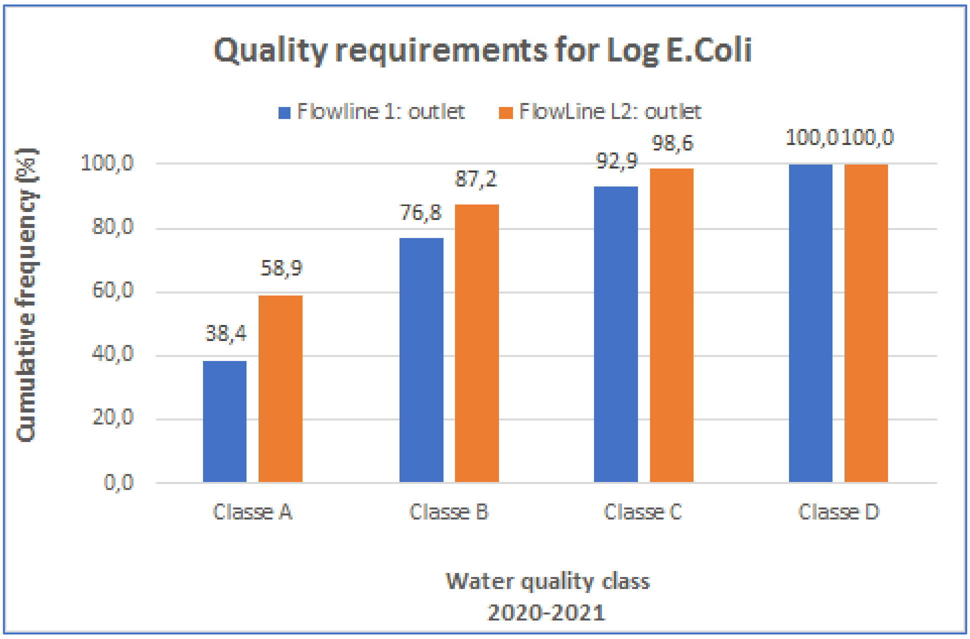


Figure 20: Cumulative frequencies from 2020 to 2021 of the quality requirements for *E. coli* achieved in each water quality class, in compliance with the recent EU Regulation on water reuse, both for Lines 1 and 2 of Peschiera Borromeo WWTP

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

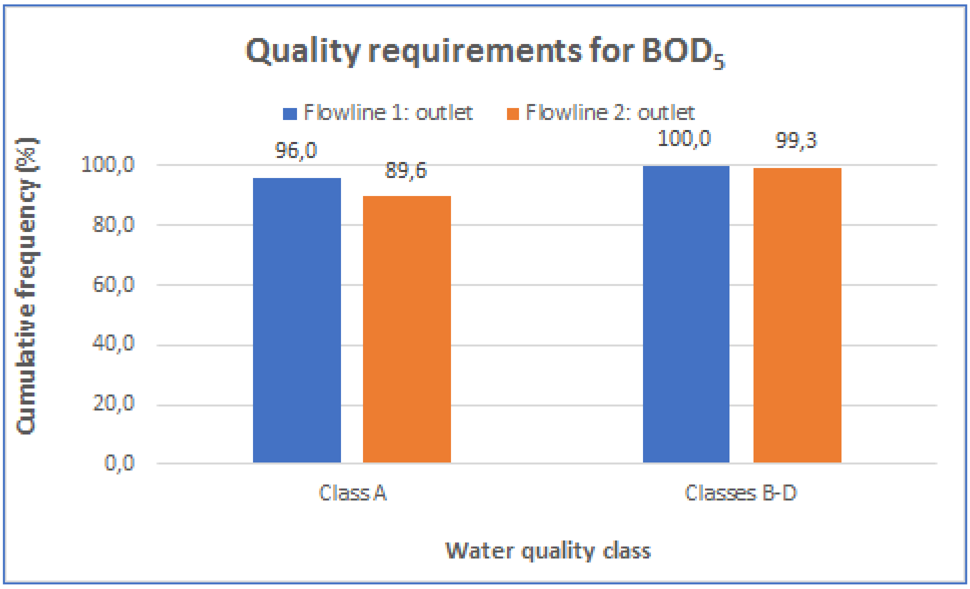


Figure 21: Cumulative frequencies from 2018 to 2021 of the quality requirements for BOD5 achieved in each water quality class, in compliance with the recent EU Regulation on water reuse, both for Lines 1 and 2 of Peschiera Borromeo WWTP

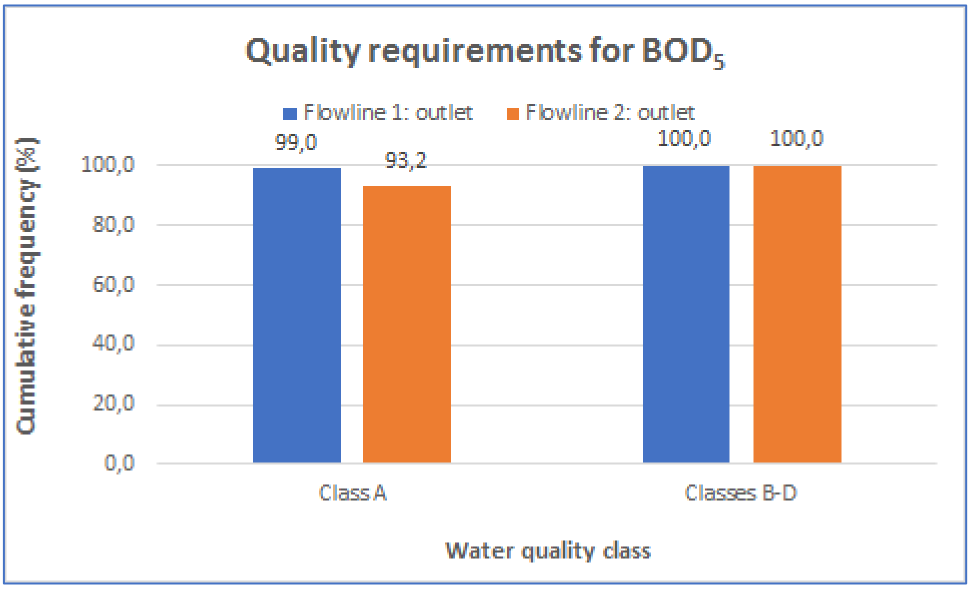


Figure 22:Cumulative frequencies from 2020 to 2021 of the quality requirements for BOD5 achieved in each water quality class, in compliance with the recent EU Regulation on water reuse, both for Lines 1 and 2 of Peschiera Borromeo WWTP

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

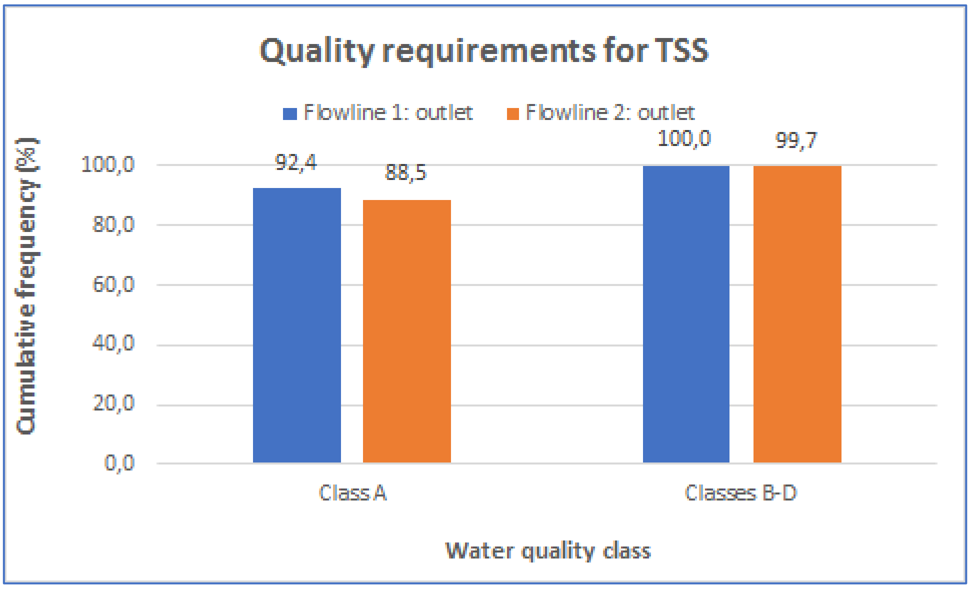


Figure 23: Cumulative frequencies from 2018 to 2021 of the quality requirements for TSS achieved in each water quality class, in compliance with the recent EU Regulation on water reuse, both for Lines 1 and 2 of Peschiera Borromeo WWTP

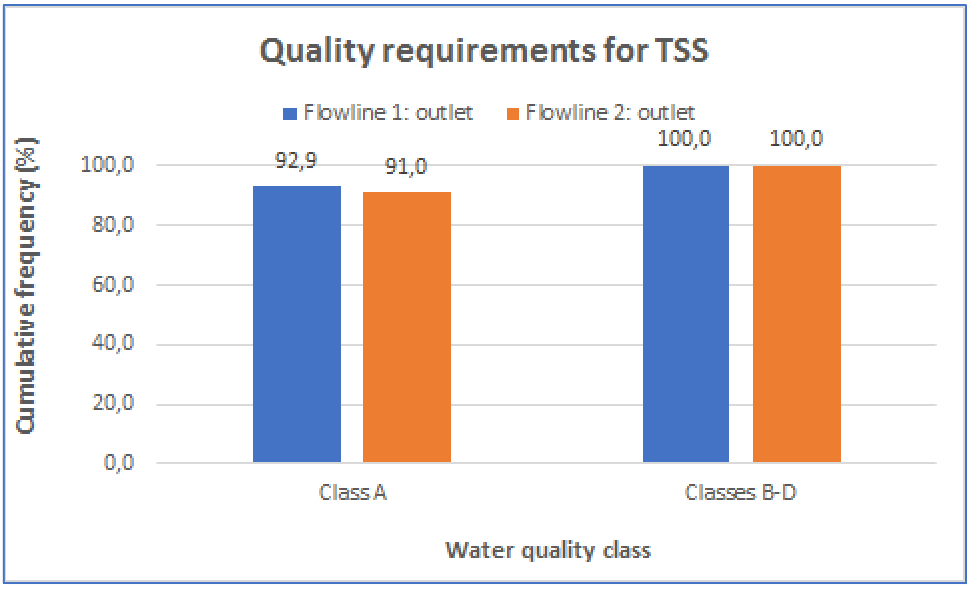


Figure 24: Cumulative frequencies from 2020 to 2021 of the quality requirements for TSS achieved in each water quality class, in compliance with the recent EU Regulation on water reuse, both for Lines 1 and 2 of Peschiera Borromeo WWTP

To assess the resilience of Peschiera Borromeo WWTP, the values of COD, total nitrogen, ammonium nitrogen, total phosphorus, BOD_5 and total suspended solids measured in the raw influent and in the final effluent of the WWTP were plotted against time, to see how much the characteristics of the treated wastewaters were affected by the incoming fluctuations. According to the graphics reported below, the inlet fluctuations (i.e., seasonal variations, occurrence of rain events, etc.) resulted well smoothed by treatments, indicating a good resilience of both lines 1 and 2. The WWTP resilience is a strategical characteristic of the plant and it has to be taken into account during the formulation of the risk matrix.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

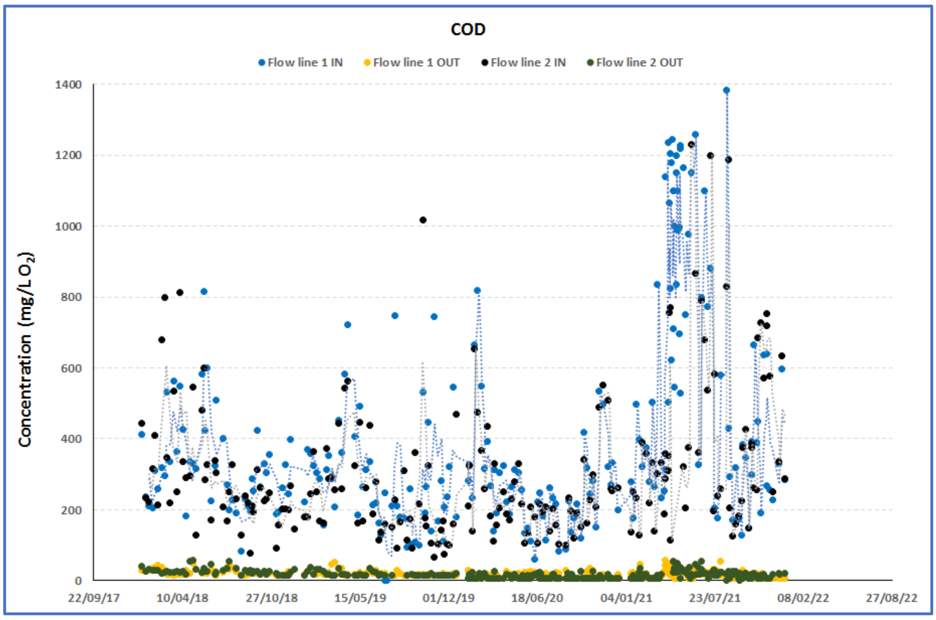


Figure 25: Influent and effluent fluctuations of COD for Lines 1 and 2 of Peschiera Borromeo WWTP from 2018 to 2021

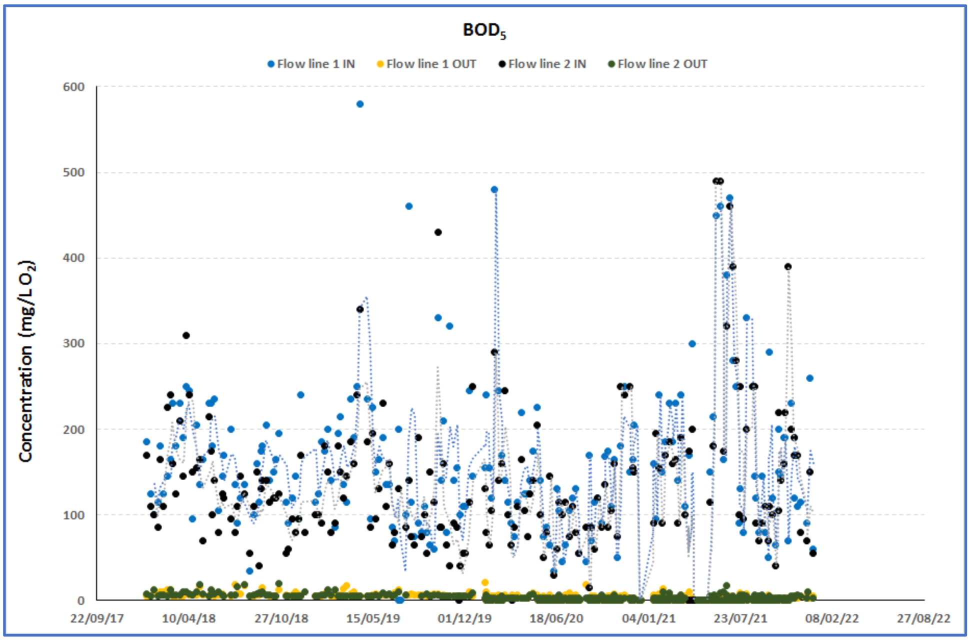


Figure 26: Influent and effluent fluctuations of BOD5 for Lines 1 and 2 of Peschiera Borromeo WWTP from 2018 to 2021

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

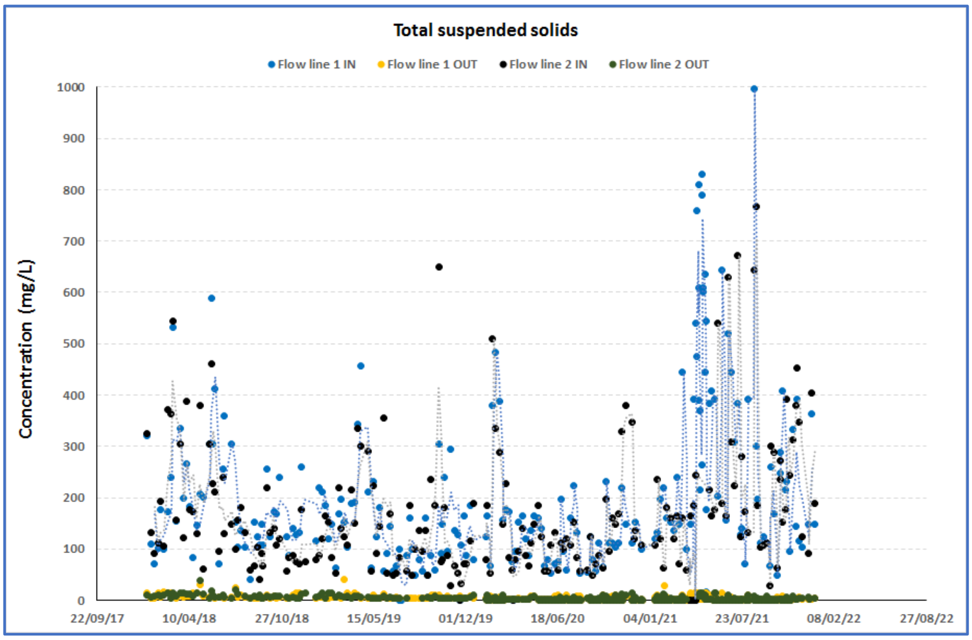


Figure 27: Influent and effluent fluctuations of Total Suspended Solids (TSS) for Lines 1 and 2 of Peschiera Borromeo WWTP from 2018 to 2021

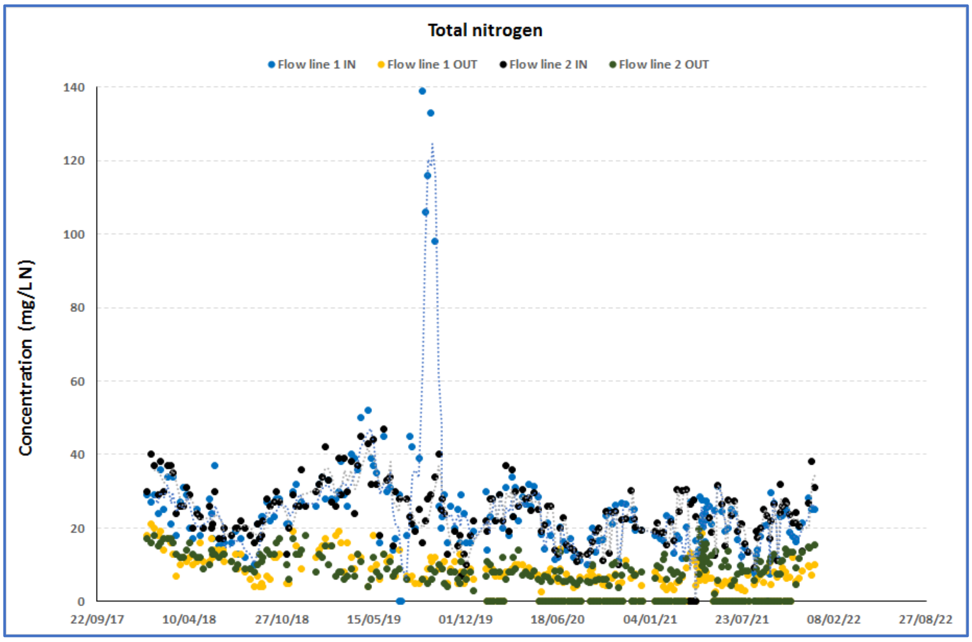


Figure 28: Influent and effluent fluctuations of Total Nitrogen for Lines 1 and 2 of Peschiera Borromeo WWTP from 2018 to 2021

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

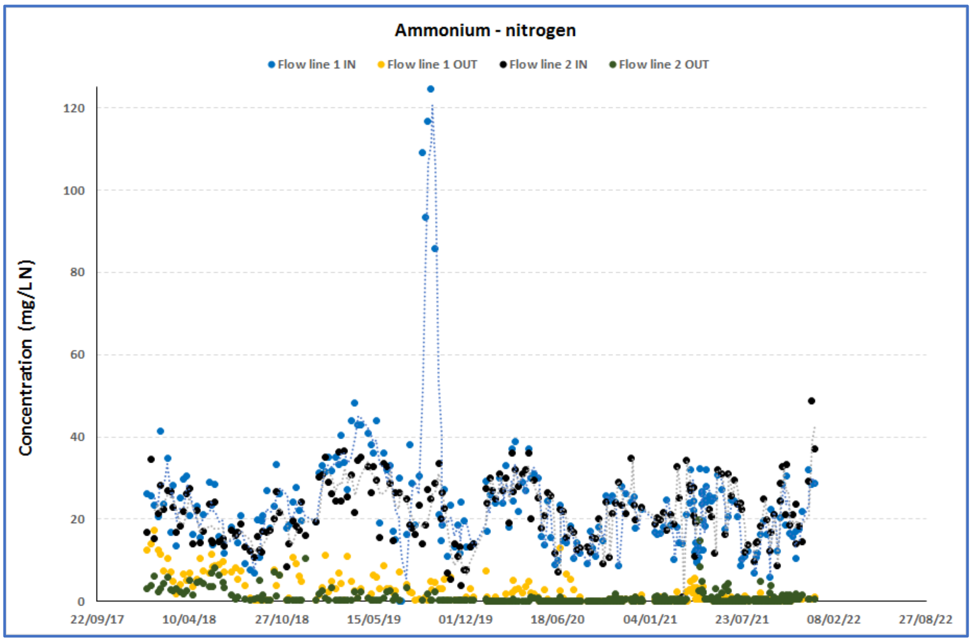


Figure 29: Influent and effluent fluctuations of Ammonium - Nitrogen for Lines 1 and 2 of Peschiera Borromeo WWTP from 2018 to 2021

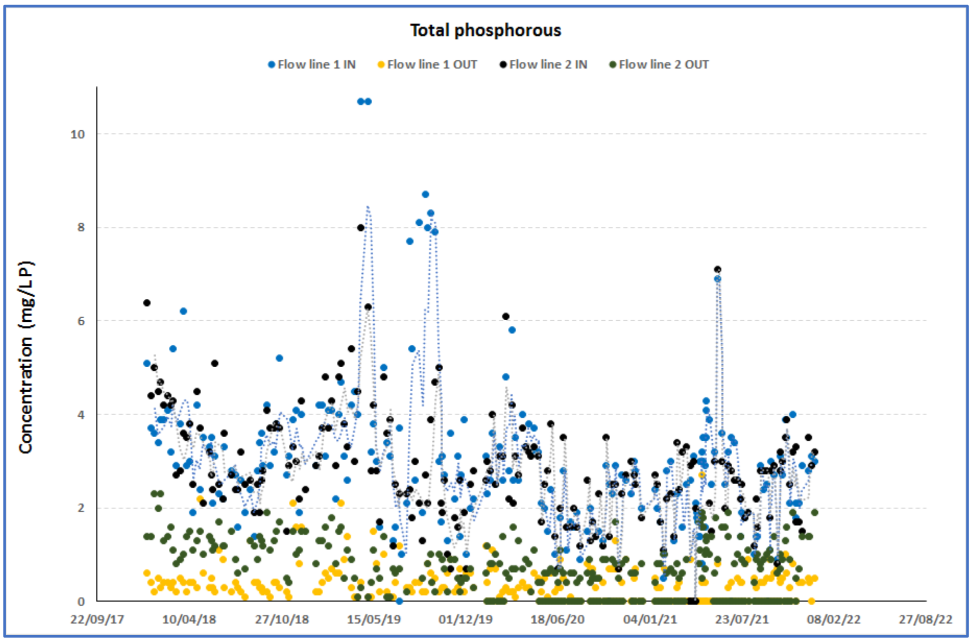


Figure 30: Influent and effluent fluctuations of Total Phosphorous for Lines 1 and 2 of Peschiera Borromeo WWTP from 2018 to 2021

3.2.3.2. Choice of treatment Line 2

The choice of Treatment Line 2 to perform water reuse is directly linked to its technical configuration: indeed, the presence of a final disinfection section based on UV irradiation gives assurance that the

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

limits for residual microbial content of the water sent to reuse could be reached with significant effectiveness.

The comparison with Treatment Line 1, which exploits chemical disinfection through peracetic acid dosage, clearly shows that this process is not able to assure adequate disinfection for water reuse under the current operational configuration and a process improvement is still needed. However, even though the chemical disinfection can be optimized to achieve comparable disinfection performances to the UV treatment, a cost comparison will lead to a significant convenience of the UV treatment. This last evidence strengthens the choice of Treatment Line 2 as preferable treatment line for the EWS implementation

3.2.4. Risk analysis (Sanitation Safety Plan) of the production chain

If the analysis of lab data gives information on WWTP performance and resilience for water reuse, risk analysis (Sanitation Safety Plan) indicates the need for an improvement of a treatment process and give indication of the possible solutions to implement to reach this scope

According to WHO SSP Manual (WHO, 2016b), three different approaches to risk assessment are available:

1. Team-based descriptive risk assessment decision

2. Semi-quantitative risk assessment, using a matrix of likelihood and severity

3. Quantitative methods (e.g., QMRA, QCRA, etc.)

To develop an Early Warning System for safe water reuse in Peschiera Borromeo, a semi-quantitative risk matrix and a Quantitative Microbiological and Chemical Risk Assessment (respectively QMRA and QCRA) were adopted for the risk analysis of the water reuse system. In Table 17, the main features of the two approaches are reported, as suggested by the Quantitative Microbial Risk Assessment guideline of WHO (WHO, 2016a).

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Table 17: Comparison of Risk matrices and QMRA (QCRA) assessment approaches

Approach	Strengths	Limitations	Resources	Expertise
Risk matrices	Approach allows for a more comprehensive consideration of hazards/hazardous events than sanitary inspections Simple prioritization structure that allows different scenarios to be compared and supports identification and management of the most important risks	Limited precision for comparing hazards or hazardous events Based on expert judgement Sometimes may be difficult to agree on risk scores; when clear and robust definitions for the likelihood and severity categories are not developed and applied, there can be inconsistent risk scoring and imbalance between acute and chronic health effects	Site visit Relevant sanitary inspection form/ checklist (Water quality assays: field or laboratory) hazardous events and severity of hazards (e.g. hazard(ous event) databases, historical records)	Interdisciplinary WSP (or SSP) team with broad qualifications covering the supply chain from source to exposure (e.g. engineers, water quality experts, catchment experts)
QMRA	Quantitative outcomes for quantitative problems Direct input from statistical inference of observational data	Most complex, requiring the most expertise and data Data are limited for quantifying model inputs Uncertainty is often difficult to quantify and incorporate in risk outcomes Validity of default assumptions may be difficult or impossible to establish for study site	All of the above plus: Quantitative data or assumptions regarding pathogen occurrence, exposure and health impacts For more in-depth QMRAs, computational tools may be required	Risk assessor(s) relying on the expertise of the WSP team and all available data Risk assessors need to be knowledgeable about interpretation of microbial data sets and, for in-depth QMRA, modelling approaches, statistics, etc.

3.2.4.1. Risk Matrix for the agricultural reuse of the wastewater treated at Peschiera Borromeo WWTP

To elaborate a semi-quantitative risk matrix, the main steps suggested by WHO Guidelines (WHO, 2016a) and reported in

Figure 31 were followed. After the description of the system, a process flow diagram (Figure 32) was drafted by the team members of ISS, CAP, UNIVPM, UNIMI.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

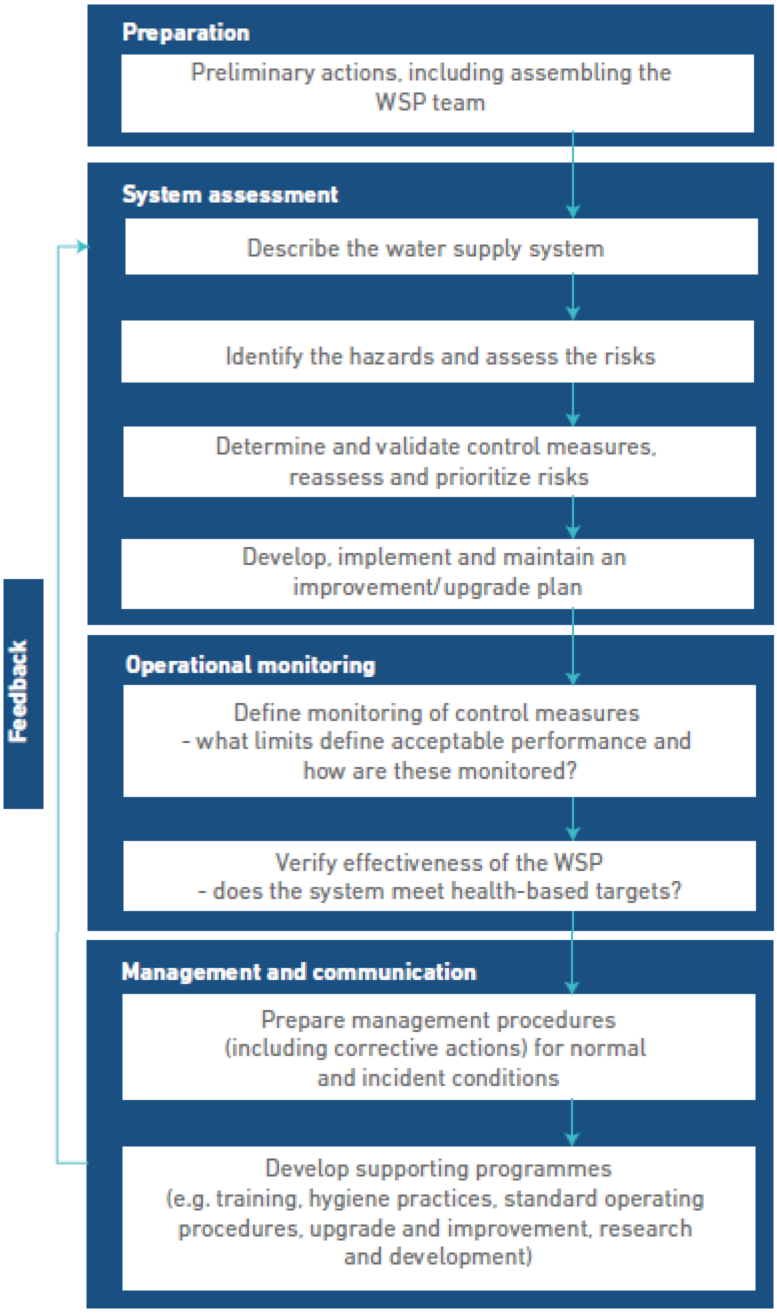


Figure 31: Steps to elaborate a semi-quantitative risk matrix within a WSP/SSP

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

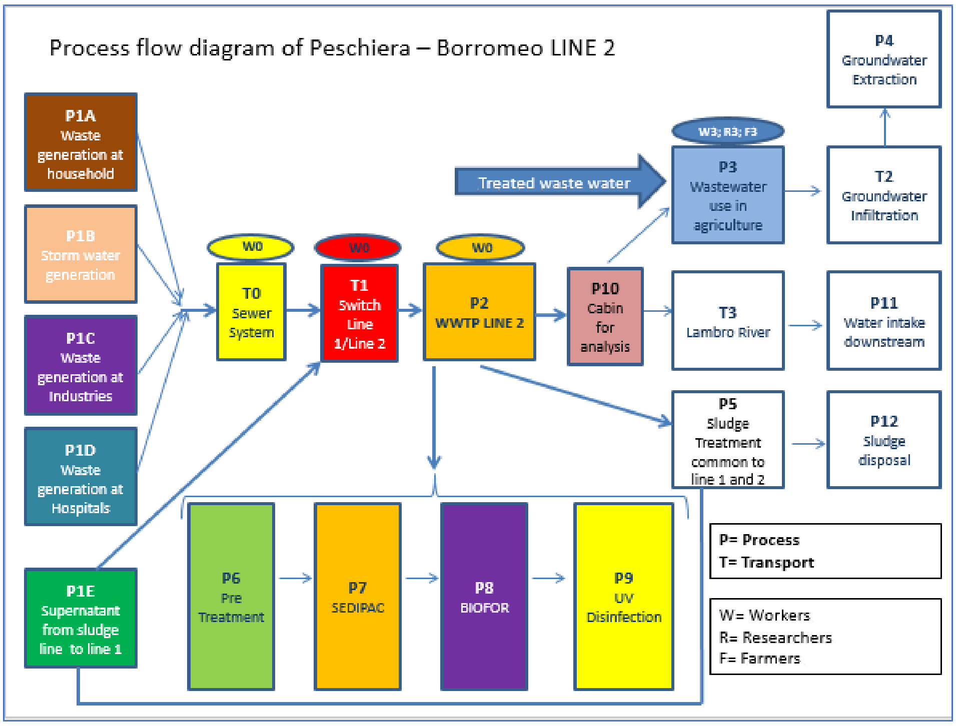


Figure 32: Process flow diagram of Peschiera Borromeo WWTP Line 2

Then, for each plant section, which is represented in Figure 32 by coloured and numbered boxes named "nodes" that are representative of process unit "P" or transport phase "T", a likelihood and a severity score were assigned for specific hazard events. These hazard events were identified by a preliminary checklist analysis (Figure 33) and by several interviews to technical operators and stakeholders of the team. The risk score R was calculated through the formula:

$$\boxed{R = L\ (\text{Likelihood}) \times S\ (\text{Severity})}$$

The scores used to quantify likelihood and severity were based on the scores reported on Figure 34, while specific descriptions were defined for Peschiera Borromeo WWTP, as reported in Table 18.

As a result, a semi-quantitative risk assessment for each identified hazardous event (i.e., microbiological, chemical, physical, radiological and reuse service interruption) was obtained and organized in matrix form.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

		NUMBER OF PARALLEL PROCESS LINES			
PRETREATMENTS	☐ Treatm. NOT present	coarse grilling	☒	several redundant parallel units in the same plant section (specify number)	-1
			☐	bypass present	-1
			☒	automatic cleaning of the grill	-1
			☒	present single with automatic cleaning (specify cleaning method)	-1
			☐	in remote control	-1
	☐ non pres.	initial lift	☒	several redundant parallel units in the same plant section (specify number)	-2
			☐	lifting with electric pumps	-1
			☒	in remote control	-1
	☐ Treatm. NOT present	fine grilling	☒	several redundant parallel units in the same plant section (specify number)	-1
			☐	bypass present	-1
			☒	automatic cleaning of the grill	-1
			☒	present single with automatic cleaning (specify cleaning method)	-1
			☒	in remote control	-1
	☐ Treatm. NOT present	sand separation	☒	sand extraction with pumps	3
			☐	sand extraction with ejector	2
			☒	rectangular tank	
			☐	circular tank	
			☒	più unità in parallelo ridondanti nella stessa sezione di impianto (specificare il numero)	-1
			☒	in remote control	-1
	☐ Treatm. NOT present	oil separation	☐	static	3
			☒	with air blowing	2
			☒	rectangular tank	
			☐	circular tank	
			☒	several redundant parallel units in the same plant section (specify number)	-1
			☒	in remote control	-1
	☐ Treatm. NOT present	primary sedimentation	☒	rectangular tank	
			☐	circular tank	-1
			☐	measurement of the level of sedimented mud	-1
			☒	in remote control	-1
			☒	several redundant parallel units in the same plant section (specify number)	-1

Figure 33: Example of a portion of a typical check list used to acquire information about the main characteristics of each treatment unit, connected equipment, its redundancy level and the related maintenance and monitoring system

The risk assessment was carried out in three stages: a first "preliminary risk assessment" (Figure 35), a second "residual risk assessment" (Figure 36) and a third "re-assessment of the risk" (Figure 37).

- In the first stage, a likelihood and a severity score were assigned to each hazardous event to obtain a preliminary risk assessment, without taking into account any control measures already present in the system. The results of this analysis represented the worst-case scenario, and they highlighted the risks to which the reclamation water production chain is potentially exposed.
- In the second stage, a "residual risk assessment" was carried out by identifying all the control measures already in place in the system and by evaluating if these measures are effective at controlling relevant risks.
- In the last third stage, namely the "re-assessment of the risk", in order to reduce residual risk scores of each possible hazardous event, supplementary control and mitigation measures were identified. Particularly, the identified mitigation and control measure are able to act on the occurrence of different hazardous events ensuring at the same time the mitigation of various risks.

It should be noted that all the mitigation measures reported in the risk matrix aim at preventing contamination events or service interruptions. Therefore, these measures act by modifying and substantially reducing the probability of occurrence of the hazardous events, with a consequent decrease of the correlated risks.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

			SEVERITY (S)				
			Insignificant	Minor	Moderate	Major	Catastrophic
			1	2	4	8	16
LIKELIHOOD (L)	Very unlikely	1	1	2	4	8	16
	Unlikely	2	2	4	8	16	32
	Possible	3	3	6	12	24	48
	Likely	4	4	8	16	32	64
	Almost Certain	5	5	10	20	40	80
Risk Score R = (L) x (S)			<6	7–12		13–32	>32
Risk level			Low Risk	Medium Risk		High Risk	Very High Risk

Figure 34: Semi-quantitative risk assessment matrix (from SSP Manual by WHO)

Table 18: Risk definitions for semi-quantitative risk assessment of Peschiera Borromeo WWTP

Descriptor		Description
Likelihood (L)		
1	Very Unlikely	Has not happened in the past and it is **highly improbable** it will happen in the next months (or another reasonable period).
2	Unlikely	Has not happened in the past but **may occur in exceptional circumstances** in the next 12 months (or another reasonable period).
3	Possible	May have happened in the past and/or **may occur under regular circumstances** in the next 12 months (or another reasonable period).
4	Likely	Has been observed in the past and/or is **likely** to occur in the next 12 months (or another reasonable period).
5	Almost Certain	Has often been observed in the past and/or **will almost certainly occur** in most circumstances in the next 12 months (or another reasonable period).
Severity (S)		
1	Insignificant	Hazard or hazardous event resulting in **no or negligible health effects** compared to background levels.
2	Minor	Hazard or hazardous event potentially resulting in **minor health effects** and/or may lead to legal complaints and concern; and/or minimal regulatory non-compliance (downgrading of the quality of the refined water of 1 class, distributed for about 1% of the time).
4	Moderate	Hazard or hazardous event potentially resulting in a **self-limiting health effects or minor illness** and/or may lead to legal complaints and concern; and/or minor regulatory non-compliance (downgrading of the quality of the refined water of 1 class, distributed for about 10% of the time).
8	Major	Hazard or hazardous event potentially resulting in **illness or injury** and/or may lead to legal complaints and concern; and/or major regulatory non-compliance (downgrading of the quality of the refined water of 2 classes).
16	Catastrophic	Hazard or hazardous event potentially resulting in **serious illness or injury, or even loss of life** and/or will lead to major investigation by regulator with prosecution likely.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

LINE 2	RISK MATRIX PESCHIERA-BORROMEO WWTP FOR AGRICULTURAL IRRIGATION WATER REUSE (Class B reclaimed water quality)												
		HAZARDS		PRELIMINARY RISK ASSESSMENT									
NODE P=Process T=Transport	WWTP SECTION	HAZARDOUS EVENT (Potential or found events able to introduce one or more specific hazards)	OBSERVATIONS (Based on bibliographic notes, operative personnel, monitoring data, inspections, users complaints)	MICROBIOLOGICAL	CHEMICAL	FISICAL	RADIOLOGICAL	REUSE SERVICE INTERRUPTION	LIKELIHOOD	SEVERITY	SCORE	RISK RATING
ALL NODES	ALL	All potential or found events able to cause damage to human health to workers (those who maintain the sewer systems, collect and transfer faecal sludge, operate the plant) through exposure routs like: ingestrion, contact, inhalation, consuption		x	x	x	x		4	16	64	Very High Risk

Figure 35: Example of a portion of matrix related to the “preliminary risk assessment”

RESIDUAL RISK					
EXISTING CONTROL MEASURES	BIBLIOGRAPHIC NOTES - CONSIDERATIONS RELATED TO THE STATISTICAL TREATMENT OF DATA (from D 1.1 e D1.3)	LIKELIHOOD	SEVERITY	SCORE	RISK RATING
• Identification of all parties involved in the water reuse system and adequate personal safety training procedures, included the use of personal protective equipment, hand washing and personal hygiene according to a clear description of their roles and responsibilities (organization chart)		1	2	2	Low risk

Figure 36: Example of a portion of matrix related to the "residual risk assessment"

RE-ASSESSMENT OF THE RISK								
INTREGRATIVE CONTROL MEASURES	VALIDATION OF CONTROL MEASURES	LIKELIHOOD	SEVERITY	SCORE	RISK RATING	OPERATIONAL MONITORING	NOTES	VERIFICATION MONITORING
• None								

Figure 37: Example of a portion of matrix related to the "re-assessment of the risk"

The entire Risk Matrix, consisting of 102 rows, is reported in Appendix A, while below are reported selected rows of the Matrix that are related to different nodes of the process flow diagram of Peschiera-Borromeo WWTP (Figure 34): *All section*, *T0 - Sewer Network*, *P1 - WWTP inlet*, *P2 - WWTP treatments*, *P3 - Treated wastewater use in agriculture*.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Moreover, since research activities of the DWC project took place during the pandemic SARS-cov-2, a special paragraph dedicated to the elements of risk analysis during "Pandemic Emergency" is also reported.

Elements of risk analysis common to all sections

Among the different elements of risk common to all sections, here it is reported the rows of the Matrix related to "Flood events". The part of preliminary risk assessment is shown in Figure 38, while the residual risk assessment and the part of the matrix related to the re-assessment of the risk are shown in Figure 40, where are also indicated possible integrative control measures to implement at Peschiera Borromeo WWTP.

LINE 2	RISK MATRIX PESCHIERA-BORROMEO WWTP FOR AGRICULTURAL IRRIGATION WATER REUSE (Class B reclaimed water quality)											
		HAZARDS		PRELIMINARY RISK ASSESSMENT								
NODE P=Process T=Transport	WWTP SECTION	HAZARDOUS EVENT (Potential or found events able to introduce one or more specific hazards)	OBSERVATIONS (Based on bibliographic notes, operative personnel, monitoring data, inspections, users complaints)	MICROBIOLOGICAL	CHEMICAL	FISICAL	RADIOLOGICAL	REUSE SERVICE INTERRUPTION	LIKELIHOOD	SEVERITY	SCORE	RISK RATING
ALL NODES	ALL SECTIONS	Flood events that cause damage to structures, contamination of the different sections of the plant as well as long periods of reuse service interruption	• Near attention zone • It seems to have occurred in the past (before 2018)	x	x	x	x	x	2	8	16	**High Risk** illness or injury and/or may lead to legal complaints and concern; and/or major regulatory non-compliance (downgrading of refined water quality by 2 classes)

Figure 38: Preliminary Risk Assessment of row 7 (first part of the matrix)

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

RESIDUAL RISK						
EXISTING CONTROL MEASURES	BIBLIOGRAPHIC NOTES - CONSIDERATIONS RELATED TO THE STATISTICAL TREATMENT OF DATA (from D 1.1 e D1.3)	LIKELIHOOD	SEVERITY	SCORE	RISK RATING	INTREGRATIVE CONTROL MEASURES
• Presence of sensor levels on equalization tanks (available for the fraction of sewer managed by CAP, serving Line 1, provided with remote control of 8 equalization tanks) • Periodic controls of the status of pumping equipment in the equalization tanks • Presence of Q influent and Q effluent from equalization tanks • Presence of level sensors and Q meters in the WWTP operative units • Presence of Q by-passed • Presence of alarms connected to level sensors and/or Q meters in the WWTP operative units • Existence of a monitoring system by radar and safety float of the level at the head of the system which, if necessary, prevent flooding by opening a gate that allows the manifold to flow into the bypass • Presence of emergency generators • Bypass maintenance • Equipment maintenance and personal training procedures • Non-compliance procedure	• From 2017 to 2021, the influent fluctuations of COD, total nitrogen, ammonium nitrogen, total phosphorus, BOD5 and total suspended solids parameters (due to seasonal variations, meteorological events, the hourly composition of the wastewater, etc.) resulted well smoothed by treatments, indicating a good resilience of the WWTP • In compliance with the EU Regulation, it resulted that the cumulative frequencies of quality requirements for E. coli, regarding the quality class B of our interest, was achieved for 69.5% of the time from 2017 to 2021 and for 87.2% of the time from 2020 to 2021	1	8	8	**Medium Risk** illness or injury and/or may lead to legal complaints and concern; and/or major regulatory non-compliance (downgrading of refined water quality by 2 classes)	• Rain heights from pluviometry data (ARPA LOMBARDIA) and warning at a defined threshold of value • Temperature and data from weather stations • DS 11: Sewer flow forecast toolbox • DS 12: Interoperable DSS and real-time control algorithms for stormwater management • DS 13: Web platform for integrated sewer and WWTP control • DS 14: Low-cost temperature sensors for real-time CSO and flooding monitoring

Figure 39: Residual Risk Assessment of row 7 (second part of the matrix)

Elements of risk analysis in T0 [Sewer Network] and T1 [Interconnection between line 1 and Line 2]

For sections T0 and T1 of the flow diagram of the wastewater reclamation system (Figure 32), are reported as examples the rows of the Risk matrix related to the occurrence of "Accidental and Illicit discharges" (Figure 40and Figure 41) and to the occurrence of "unexpected load in Line 2" (Figure 42, Figure 43). For both the hazardous events different DSS are listed among the integrative control measures to reduce the risk.

RISK MATRIX PESCHIERA-BORROMEO WWTP FOR AGRICULTURAL IRRIGATION WATER REUSE (Class B reclaimed water quality)											
	HAZARDS	PRELIMINARY RISK ASSESSMENT									
WWTP SECTION	HAZARDOUS EVENT (Potential or found events able to introduce one or more specific hazards)	OBSERVATIONS (Based on bibliographic notes, operative personnel, monitoring data, inspections, users complaints)	MICROBIOLOGICAL	CHEMICAL	FISICAL	RADIOLOGICAL	REUSE SERVICE INTERRUPTION	LIKELIHOOD	SEVERITY	SCORE	RISK RATING
SEWER NETWORK (52 industrial discharges DLGs. 152/06 - 23% of total load)	**Accidental or illicit discharges (es.TIR accidental overflows, discharges of solvents, hydrocarbons, ammonia, etc.)**	• Verified in the past, monitored by national environmental agency • Verified in the past illicit discharges at night-time • At the moment there is no online monitoring of the inlet	x	x	x	x	x	4	8	32	**High Risk** illness or injury and/or may lead to legal complaints and concern; and/or major regulatory non-compliance (downgrading of refined water quality by 2 classes)

Figure 40: Preliminary Risk Assessment of row 21 (first part of the matrix)

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

RESIDUAL RISK						
EXISTING CONTROL MEASURES	BIBLIOGRAPHIC NOTES - CONSIDERATIONS RELATED TO THE STATISTICAL TREATMENT OF DATA (from D 1.1 e D1.3)	LIKELIHOOD	SEVERITY	SCORE	RISK RATING	INTREGRATIVE CONTROL MEASURES
• P2 Treatments • Warnings from Environmental Protection Agency control activity on illicit or accidental discharges • Presence of outlet probes • Equipment maintenance and personal training procedures • Non-compliance procedure • Emergency procedure	• From 2017 to 2021, the influent fluctuations of COD, total nitrogen, ammonium nitrogen, total phosphorus, BOD5 and total suspended solids parameters (due to seasonal variations, meteorological events, the hourly composition of the wastewater, etc.) resulted well smoothed by treatments, indicating a good resilience of the WWTP **• In compliance with the EU Regulation, it resulted that the cumulative frequencies of quality requirements for E. coli, regarding the quality class B of our interest, was achieved for 69.5% of the time from 2017 to 2021 and for 87.2% of the time from 2020 to 2021**	1	8	8	**Medium Risk** illness or injury and/or may lead to legal complaints and concern; and/or major regulatory non-compliance (downgrading of refined water quality by 2 classes)	**• Integration of DS: 4, 8, 9, 11, 12,13, 14, 15 (in particular DS 8: DTS sensor for tracking illicit sewer connections and DS 9: Sensors and smart analytics for tracking illicit sewer connections hotspots)** **• Control room input and output sensor data management to effectively modulate treatments** **• In the medium to long term, building of water reservoirs to avoid reuse service interruption**

Figure 41: Residual Risk Assessment of row 21 (second part of the matrix)

LINE 2	RISK MATRIX PESCHIERA-BORROMEO WWTP FOR AGRICULTURAL IRRIGATION WATER REUSE (Class B reclaimed water quality)												
		HAZARDS	PRELIMINARY RISK ASSESSMENT										
NODE P=Process T=Transport	WWTP SECTION	HAZARDOUS EVENT (Potential or found events able to introduce one or more specific hazards)	OBSERVATIONS (Based on bibliographic notes, operative personnel, monitoring data, inspections, users complaints)	MICROBIOLOGICAL	CHEMICAL	FISICAL	RADIOLOGICAL	REUSE SERVICE INTERRUPTION	LIKELIHOOD	SEVERITY	SCORE	RISK RATING
T1	INTERCONNECTION BETWEEN LINE1-LINE2	**Unexpected loads at Line 2, usually managed in Line 1**	• Sometimes Line 2 could serve flows usually tretaed in Line 1 (92 industries served in Line 1) • There is an internal barrier, always open to allow the access at the bypass • At Line 1 is collected the centrate from sludge line, rich in N. The centrate could arrive in Line 2, after dilution with sewage flow	x	x	x	x	x	4	8	32	**High Risk** illness or injury and/or may lead to legal complaints and concern; and/or major regulatory non-compliance (downgrading of refined water quality by 2 classes)

Figure 42: Preliminary Risk Assessment of row 23 (first part of the matrix)

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

RESIDUAL RISK						
EXISTING CONTROL MEASURES	BIBLIOGRAPHIC NOTES - CONSIDERATIONS RELATED TO THE STATISTICAL TREATMENT OF DATA (from D 1.1 e D1.3)	LIKELIHOOD	SEVERITY	SCORE	RISK RATING	INTREGRATIVE CONTROL MEASURES
• P2 Treatments • Presence of level sensor on overflow section	• From 2017 to 2021, the influent fluctuations of COD, total nitrogen, ammonium nitrogen, total phosphorus, BOD5 and total suspended solids parameters (due to seasonal variations, meteorological events, the hourly composition of the wastewater, etc.) resulted well smoothed by treatments, indicating a good resilience of the WWTP • **In compliance with the EU Regulation, it resulted that the cumulative frequencies of quality requirements for E. coli, regarding the quality class B of our interest, was achieved for 69.5% of the time from 2017 to 2021 and for 87.2% of the time from 2020 to 2021**	1	8	8	**Medium Risk** illness or injury and/or may lead to legal complaints and concern; and/or major regulatory non-compliance (downgrading of refined water quality by 2 classes)	• Integration of DS: 3, 8, 9, 11, 12,13, 14, 15, 4 • **Control room input and output sensor data management to effectively modulate treatments** • **Building of a new treatment line (LINE 3)**

Figure 43: Residual Risk Assessment of row 23 (second part of the matrix)

Elements of risk analysis in P1 - WWTP inlet

The most important element of risk in P1 section is the entrance to the WWTP of hazardous pathogens (Table 19) and hazardous substances (tab 3/A, Part III all.5 Italian Legislative Decree 152/06).

Table 19: Concentration of microbial contaminant in wastewater

	RAW WASTEWATER	CONCENTRATION		REFERENCE
bacteria	*Total coliform*	CFU/100ml	3.9E+07	(Kay et al., 2008)
		CFU/100ml	1.96e+07 - 4.36E+07	(WERF, 2004)
		CFU/100ml	5.15e+08	(Thwaites et al., 2018)
		MPN/100ml	10^7 - 10^9	(Metcalf & Eddy, 2014)
		MPN/100ml	1.1E+08	(Howard et al., 2004)
	Fecal coliforms	MPN/100ml	10^6 - 10^8	(Metcalf & Eddy, 2014)
		MPN/100ml	8.2E+06	(Howard et al., 2004)
		MPN/100ml	2.1E+06 - 5.3E+06	(Oakley & Mihelcic, 2019)
		CFU/100ml	1.7E+07	(Kay et al., 2008)
		CFU/100ml	2.09e+06 - 5.31e+06	(WERF, 2004)
	E.coli	MPN/100ml	10^5 - 10^7	(Metcalf & Eddy, 2014)
		MPN/100ml	5.37E+03 - 3.47E+07	(Oakley & Mihelcic, 2019)
		logMPN/100 ml	6.42±0.28 (5.99 - 7.28)	(Bailey et al., 2018)
		CFU/100ml	3.6e+06	(Marín et al., 2015)
		CFU/100ml	5.31e+07	(Thwaites et al., 2018)

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

		N/l	10^5 - 10^10	(NRMMC-EPHC-AHMC, 2006)
	Salmonella	MPN/100ml	10^2 - 10^4	(Metcalf & Eddy, 2014)
		MPN/100ml	266.7	(Howard et al., 2004)
		MPN/L	3-1100	(Lemarchand & Lebaron, 2003)
		logMPN/100 ml	4.08±0.99 (2.88 - 5.88)	(Bailey et al., 2018)
		n/l	10^3 - 10^5	(NRMMC-EPHC-AHMC, 2006)
		n/l	1-10^5	(WHO, 2006)
		N/l	up to 10^5	(US EPA, 2012)
protozoa parasites	*Cryptosporidium*	MPN/100ml	10 - 10^3	(Metcalf & Eddy, 2014)
		n/l	1-10^4	(WHO, 2006)
		n/l	up to 10^4	(US EPA, 2012)
		oocysts/L	0 - 10^4	(NRMMC-EPHC-AHMC, 2006)
		oocysts/L	2.5 - 277	(Hamilton et al., 2018)
		oocysts/L	4.5±0.8	(Carraro et al., 2000)
		oocysts/L	1-87.13	(Lemarchand & Lebaron, 2003)
		oocysts/L	96±105	(Ramo et al., 2017)
		oocysts/L	6–350	Montemayor et al., 2005; Castro-Hermida et al., 2008, 2010; Galván et al., 2014
		oocysts/L	22 - 456	(Oakley & Mihelcic, 2019)
		oocysts/L	6.5 - 37.8	(Oakley & Mihelcic, 2019)
		oocysts/L	6.55 - 37.8	(WERF, 2004)
		oocysts/L	0.3 - 5e+04	(Soller et al., 2017)
		oocysts/L	30±5	(Briancesco & Bonadonna, 2005b)
		oocysts/L	371±24 (64.6 - 955)	(Bailey et al., 2018)
	Giardia	MPN/100ml	10^3 - 10^4	(Metcalf & Eddy, 2014)
		N/L	10^2 - 10^5	(WHO, 2006)
		N/L	up to 10^5	(US EPA, 2012)
		cysts/L	10^2 - 10^5	(NRMMC-EPHC-AHMC, 2006)
		cysts/L	764 - 6606	(Oakley & Mihelcic, 2019)
		cysts/L	3247±2039	(Ramo et al., 2017)
		cysts/L	61.2 - 794	(Oakley & Mihelcic, 2019)
		cysts/L	61.2 - 794	(WERF, 2004)

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

		cysts/L	89-8305	Montemayor et al., 2005; Castro-Hermida et al., 2008, 2010; Galván et al., 2014
		cysts/L	3.2 - 1.0E+04	(Soller et al., 2017)
		cysts/L	7000 ± 2000	Briancesco et al. 2004
		cysts/L	302 ± 24 (45 - 1122)	(Bailey et al., 2018)
viruses	*Enterovirus*	MPN/100ml	10^3 - 10^4	(Metcalf & Eddy, 2014)
		MPN/100ml	4.64E+02 - 9.37E+03	(WERF, 2004)
		MPN/L	4.6 - 93.7	(Oakley & Mihelcic, 2019)
		N/L	10^5 - 10^6	(WHO, 2006)
		N/L	up to 10^6	(US EPA, 2012)
		PFU/L	10^2 - 10^6	(NRMMC-EPHC-AHMC, 2006)
	Norovirus	PFU/L	10 - 10^4	(NRMMC-EPHC-AHMC, 2006)
		copies/L	10^3.76 ±10^0.93	(Soller et al., 2017)
	Norovirus GI	copies/L	1.12 E+03 - 5.75E+05	(Oakley & Mihelcic, 2019)
	Norovirus GII	copies/L	6.46e+02 - 2.19e+06	(Oakley & Mihelcic, 2019)

Generally, the wastewater coming to a WWTP is composed by different streams, which include domestic wastewaters (P1A), stormwater (P1B), industrial wastewater (P1C), hospital wastewater (P1D) and supernatant from sludge line (P1E). Different pathogens are present in these streams, which are then mixed together before entering the WWTP. The parts of the matrix related to the presence of hazardous pathogens in the influent wastewater are reported in Figure 44 and Figure 45.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

LINE 2	RISK MATRIX PESCHIERA-BORROMEO WWTP FOR AGRICULTURAL IRRIGATION WATER REUSE (Class B reclaimed water quality)											
		HAZARDS		PRELIMINARY RISK ASSESSMENT								
NODE P=Process T=Transport	WWTP SECTION	HAZARDOUS EVENT (Potential or found events able to introduce one or more specific hazards)	OBSERVATIONS (Based on bibliographic notes, operative personnel, monitoring data, inspections, users complaints)	MICROBIOLOGICAL	CHEMICAL	FISICAL	RADIOLOGICAL	REUSE SERVICE INTERRUPTION	LIKELIHOOD	SEVERITY	SCORE	RISK RATING
P1	WWTP INLET P1A Domestic Wastewaters P1B Stormwater P1C Industrial Wastewater P1D Hospital Wastewater P1E Supernatant from sludge line to Line 1	Hazardous pathogens (reported in literature) and hazardous substances (tab 3/A, Part III all.5 D.Lgs. 152/06) in the influent	• Anomalous discharges have been detected (solvents, glues) that affect treatment efficiency • Lab analyses and flows are registered in the daily register of the plant "Registro Giornaliero di Funzionamento Impianto(RGFI)" • Lack of sufficient information to carry out a risk assessment for each subsection P1A, P1B, P1C, P1D, P1E • Literature data on pathogens occurrence in wastewater	x	x	x	x	x	5	8	40	**Very High** Hazard or hazardous event potentially resulting in serious illness or injury, or even loss of life and/or will lead to major investigation by regulator with prosecution likely

Figure 44: Preliminary Risk Assessment of row 31 (first part of the matrix)

RESIDUAL RISK						
EXISTING CONTROL MEASURES	BIBLIOGRAPHIC NOTES - CONSIDERATIONS RELATED TO THE STATISTICAL TREATMENT OF DATA (from D 1.1 e D1.3)	LIKELIHOOD	SEVERITY	SCORE	RISK RATING	INTREGRATIVE CONTROL MEASURES
• P2 Treatments • 24h weekly sampling (after the fine screen) for laboratory analysis (frequency 3-6 times/ month) of parameters required for the minimum quality of reclaimed water class B • Presence of outlet probes • Equipment maintenance and personal training procedures • Maintenance contracts with companies • Environmental Protection Agency controls • Non-compliance procedure • Emergency procedure	• From 2017 to 2021, the influent fluctuations of COD, total nitrogen, ammonium nitrogen, total phosphorus, BOD5 and total suspended solids parameters (due to seasonal variations, meteorological events, the hourly composition of the wastewater, etc.) resulted well smoothed by treatments, indicating a good resilience of the WWTP • In compliance with the EU Regulation, it resulted that the cumulative frequencies of quality requirements for E. coli, regarding the quality class B of our interest, was achieved for 69.5% of the time from 2017 to 2021 and for 87.2% of the time from 2020 to 2021	1	8	8	**Medium Risk** illness or injury and/or may lead to legal complaints and concern; and/or major regulatory non-compliance (downgrading of refined water quality by 2 classes)	• Integration of DSS: 3, 4, 8, 9, 11, 12,13, 14, 15 • Control room input and output sensor data management to effectively modulate treatments • Upgrading of the disinfection system • In the medium to long term, building of water reservoirs to avoid reuse service interruption • Periodic samplings and analysis on industrial discharges • Periodic samplings and analysis on wastewaters from S. Raffaele Hospital • Periodic investigation on toxic, persistent and /or emerging non - regulated compounds (e.g. heavy metals and persistent organic contaminants, pharmaceuticals and personal care products, pesticides and herbicides, industrial chemicals)

Figure 45: Residual Risk Assessment of row 31 (second part of the matrix)

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Elements of risk analysis in P2 - WWTP treatments train

One of the most important elements of risk in P2 section, that represents the treatments train of Line 2 at Peschiera-Borromeo WWTP (Figure 10) is the *"Malfunctioning of UV disinfection system"* (i.e., P9 subsection in Figure 32). The related rows of the Risk Matrix are reported in Figure 47 and 48.

Concerning the UV disinfection unit, a specific literature review was conducted to find information about expected pathogens concentrations in the wastewater before and after the treatment along with the observed removal efficiencies (Table 20). It can be observed that in most of the literature studies, Log removals for pathogens during UV treatment are in the range 2-4. At Peschiera-Borromeo WWTP, after the revamping of the plant, *E. coli* concentrations in the final effluent is often lower than 10 CFU/100ml (58.8% of the measurements in 2020-2021), and thus in compliance with the limit for the quality class A of EU regulation on water reuse.

Table 20: Pathogens removals observed during UV treatment

	MICRORGANISM		PRE UV	POST UV	LOG REMOVAL	REF.
bacteria	*Bacteria*				2 - > 4	(WHO, 2006)
	E. coli				2 - > 4	(NRMMC-EPHC-AHMC, 2006)
					1.8 - 4.7	(DEMOWARE project, 2013)
		CFU/ 100ml	860 - 24000	1 - 2	2.63 - 4.38	(Francy et al., 2012)
		CFU/ml	(31±6)-(82±16)	0 - 14±2		(Anastasi et al., 2013)
	Salmonella	MPN/ 100ml	45			(Howard et al., 2004)
					5.6	(Hijnen et al., 2006)
		CFU/L		6		(Soller et al., 2017)
protozoa parasites	*Cryptosporidium*				> 3	(WHO, 2006)
					> 3	(NRMMC-EPHC-AHMC, 2006)
					3	(Chahal et al., 2016)
		N/l	6 - 23	2 - 8	0.45 - 0.48	(Liberti et al., 2003)
		oocysts/L	-	3.6		(Deng et al., 2019)
					2 - 3.5	(Soller et al., 2017)
virus	*Virus*				1 - > 3	(WHO, 2006)
					2.9 - 4.2	(DEMOWARE project, 2013)
					4	(Chahal et al., 2016)

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

	Enterovirus				> 3	(NRMMC-EPHC-AHMC, 2006)
		Log copies/L	(3.71±0.43)-(4.13±0.31)	(3.47±0.49) - (3.82±0.32)		(Qiu et al., 2018)
		copies/l	6.8 - 250	36 - 67	-	(Francy et al., 2012)
	Norovirus				0.5 - 1.5	(Soller et al., 2017)
		RT-qPCR (GC/mL)	-	3.30-3.80		(Deng et al., 2019)
	Norovirus GI	Log copies/L	0.84 - 2.4	1.56 - 1.83	-	(Francy et al., 2012)
		Log copies/L	(4.19±0.49)-(4.35±0.53)	(4.02±0.52) - (4.18±0.6)		(Qiu et al., 2018)
	Norovirus GII	Log copies/L	(4.96±0.36)-(5.23±0.29)	(4.86±0.36) - (5.12±0.31)		(Qiu et al., 2018)

According to the studies reported in the paragraph *Analysis and assessment of the WWTP efficiency and resilience*, it was observed that the cumulative frequency of quality requirements for *E. coli* in class B was achieved for 87.2% of the measurements in the period 2020 - 2021. To increase this percentage and reduce the risk to have high concentration of pathogens in the final effluent, integrative control measures were reported in the Risk Matrix (Figure 46).

LINE 2	RISK MATRIX PESCHIERA-BORROMEO WWTP FOR AGRICULTURAL IRRIGATION WATER REUSE (Class B reclaimed water quality)											
		HAZARDS	PRELIMINARY RISK ASSESSMENT									
NODE P=Process T=Transport	WWTP SECTION	HAZARDOUS EVENT (Potential or found events able to introduce one or more specific hazards)	OBSERVATIONS (Based on bibliographic notes, operative personnel, monitoring data, inspections, users complaints)	MICROBIOLOGICAL	CHEMICAL	FISICAL	RADIOLOGICAL	REUSE SERVICE INTERRUPTION	LIKELIHOOD	SEVERITY	SCORE	RISK RATING
P9 -UV DISINFECTION	**UV DISINFECTION** 2 lines of 6 tanks (3 tanks each line), divided in 114 modules, with a total of 912 lamps with a mono light lenght of 254 nm	Malfunctioning of UV disinfection treatment	• Informations about expected pathogens concentrations in the wastewater before and after the treatment along with the observed removal efficiencies are reported in 3.2.6.1.5 paragraph	x				x	3	16	48	**Very High** serious illness or injury, or even loss of life and/or will lead to major investigation by regulator with prosecution likely

Figure 46: Preliminary Risk Assessment of row 74 (first part of the matrix)

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

RESIDUAL RISK						
EXISTING CONTROL MEASURES	BIBLIOGRAPHIC NOTES - CONSIDERATIONS RELATED TO THE STATISTICAL TREATMENT OF DATA (from D 1.1 e D1.3)	LIKELIHOOD	SEVERITY	SCORE	RISK RATING	INTREGRATIVE CONTROL MEASURES
• P6, P7, P8 Treatments • The number of modules is overdimensioned • Planned maintenance program from the tech-provider • Weekly instant sampling and 24h weekly authomatic sampling for laboratory analysis (frequency 3-6 times/ month) of parameters required for the minimum quality of reclaimed water class B • Daily visual check of the electromechanical state • Check device state and alarms (request for intervention) by remote control • Equipment maintenance and personal training procedures • Non-compliance procedure • Emergency procedure	• In compliance with the EU Regulation, it resulted that the cumulative frequencies of quality requirements for E. coli, regarding the quality class B of our interest, was achieved for 69.5% of the time from 2017 to 2021 and for 87.2% of the time from 2020 to 2021	3	4	12	**Medium** May lead to legal complaints and concern; and/or minor regulatory non-compliance (downgrading of refined water quality by 1 class for about 10% of the time).	• Presence of UV dose working sensor, based on the Q effluent flow • Daily and random monitoring of lamps intensity • Evaluation of the fouling kinetics of lamps for their cleaning • Replacement of wavelenght fixed lamps at 254 nm with multi wavelenght lamps • Integrable Memo Maintenance Program in TLC • In the medium to long term, building of water reservoirs to avoid reuse service interruption • DS 1: Sensors for real-time in-situ E.coli and enterococci measurements at the effluent • DS 1.3 Early warning system for safe water reuse

Figure 47: Residual Risk Assessment of row 74 (second part of the matrix)

Elements of risk analysis in P3 – Use of treated wastewater in agriculture

As already mentioned in Section *Agricultural practices*, the pilot field in DWC project was cultivated with maize (for animal fodder and for energy production) in spring-summer and mustard (as cover crop) during winter. These crops were irrigated by Class B treated water using drip irrigation. The location of the experimental field is shown in Figure 48.

Figure 48: Position of the experimental fields in relation to Peschiera Borromeo treatment plant

One of the most insidious elements of risk in this context is the *"Uncontrolled microbial regrowth due to discontinuous use (temporary and ad-hoc use) of class B treated wastewater"* (Figure 49). Foreseen integrative control measures to avoid pathogens regrowth along the

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

irrigation infrastructures due to intermittent water supply include procedures of cleaning, washing, sanitizing, etc. during the reuse service interruption.

LINE 2	RISK MATRIX PESCHIERA-BORROMEO WWTP FOR AGRICULTURAL IRRIGATION WATER REUSE (Class B reclaimed water quality)												
		HAZARDS	PRELIMINARY RISK ASSESSMENT										
NODE P=Process T=Transport	WWTP SECTION	HAZARDOUS EVENT (Potential or found events able to introduce one or more specific hazards)	OBSERVATIONS (Based on bibliographic notes, operative personnel, monitoring data, inspections, users complaints)	MICROBIOLOGICAL	CHEMICAL	FISICAL	RADIOLOGICAL	REUSE SERVICE INTERRUPTION	LIKELIHOOD	SEVERITY	SCORE	RISK RATING
P3 - Treated wastewater use in agricolture	**ALL SUB-SECTIONS (the supply, the distribution infrastructure, the pilot field)**	**Uncontrolled microbial regrowth due to discontinuous use (temporary and ad-hoc use) of class B treated wastewater, when the supply of treated water is upon request (especially during the warm season)**	• Continuous use of Class B treated water from April to May and discontinuation use from the end of September to the beginning of March • During non-supply, stagnant class B treated water in P3 could alter the quality of the next requested supply of class B treated water (for example Legionella) • Risk of high biofilm and algae production • Risk of obstruction of irrigation pipe outlet holes	x	x	x	x	x	5	8	40	**Very High Risk** illness or injury and/or may lead to legal complaints and concern; and/or major regulatory non-compliance (downgrading of refined water quality by 2 classes).

Figure 49: Preliminary Risk Assessment of row 85 (first part of the matrix)

RESIDUAL RISK						
EXISTING CONTROL MEASURES	BIBLIOGRAPHIC NOTES - CONSIDERATIONS RELATED TO THE STATISTICAL TREATMENT OF DATA (from D 1.1 e D1.3)	LIKELIHOOD	SEVERITY	SCORE	RISK RATING	INTREGRATIVE CONTROL MEASURES
• Cultivation of the crop by drip irrigation		1	8	8	**Medium Risk** illness or injury and/or may lead to legal complaints and concern; and/or major regulatory non-compliance (downgrading of refined water quality by 2 classes)	• Drafting of adequate management procedures of P3 (cleaning, washing, sanitizing, etc.) during non-supply period, also as a function of temperature variations • Equipment maintenance and personal training procedures • USEFUL: an underground irrigation system wolud be able to avoid strong thermal excursions and microbial growth • DS 1.3 Early warning system for safe water reuse

Figure 50: Residual Risk Assessment of row 85 (second part of the matrix)

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Elements of risk analysis during “Pandemic Emergency”

During the recent pandemic emergency, several relationships and risks related to the SARS-CoV-2 virus, responsible for the cases of COVID-19 (Coronavirus Disease), were found in connection with water and sanitation infrastructures (Gruppo di lavoro ISS Ambiente Rifiuti, 2020). Considering the impressive number of infections by SARS-CoV-2, it results highly important to investigate and identify all the potential routes of transmission of the virus.

Viruses are responsible for a wide range of diseases, such as gastroenteritis, upper and lower respiratory tract syndromes, conjunctivitis, hepatitis, central nervous system infections, cardio-circulatory system infections, and chronic diseases. From the excretion of viruses with faeces, urine and other body secretions, sometimes in high concentrations, it follows that in urban wastewater and, consequently, along the integrated water cycle they can be detected in high concentrations.

Figure 51 illustrates the fate of viruses in the integrated water cycle and the stages in which potential contact with the virus by exposed individuals can occur. Possible contact pathways are also listed below:

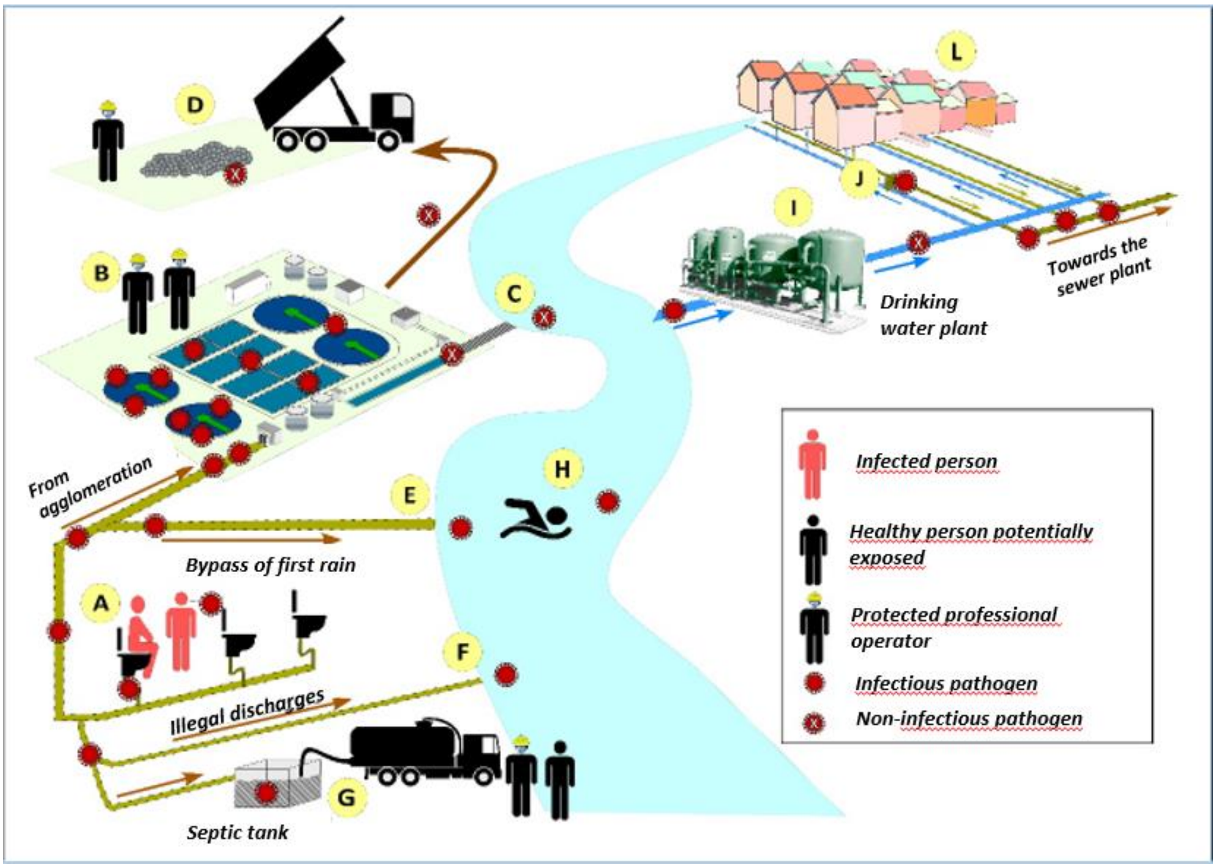


Figure 51: Fate of viral pathogens in the integrated water cycle and potential human exposure points (modified by: Wigginton et al. Environ Sci Water Res Technol 2015;1:735-46)

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

A) Viruses excreted with faeces, urine, vomiting, saliva or respiratory secretions enter the sewage system. Indoor water discharges to buildings can generate virus-laden aerosols leading to a risk of exposure.
B) Viruses are transported through the sewerage system to the wastewater treatment plant, where exposure through aerosols is limited to professional operators adequately protected by personal protective equipment (PPE).
C) Viruses entering the purification plant are generally inactivated by physical, biological and chemical treatment processes.
D) Wastewater treatment generates biosolids, that is sewage sludge that can be disposed of through spreading on land, incineration or landfilling. The exposure during sludge management and handling is limited to professional operators protected. Virus inactivation control on disposed sludge is achieved through regulatory measures and good practices as described in the ISS COVID-Report19 n. 9/2020 Interim indications on the management of sewage sludge for the prevention of the spread of SARS-cov-2 virus.
E) First rainwater generated by intense meteoric events coming from the sewerage system is channelled through a drain or bypass well, directly to the receiving water body, carrying potentially infectious viruses.
F) Illicit discharges can cause potentially contaminated wastewater to flow directly into the receiving water body.
G) Conventional biological tanks, used in the case of buildings not connected to a sewerage system, may contain viral pathogens with consequent exposure risks for the purge service operators and any subjects present in the vicinity of the places of operation.
H) Recreational activities may lead to exposure to infectious viruses present in surface water when carried upstream of the same water.
I) Water intended for human consumption undergoes a series of physico-chemical treatment processes to remove contaminants, including viruses potentially present in the collection. The risk analysis carried out according to the PSA model thoroughly examines in prevention the dangerous events and the dangers that can occur at each stage of the hydro-potable supply chain.
J) Breaks or interruptions in the sewerage system may cause contamination of drinking water if the distribution and sewerage networks come into contact; similar risks may occur inside buildings where inadequate installations or operations lead to the entry of wastewater or aerosols generated by wastewater, into water or aeration networks.
K) Users connected to the drinking water distribution network may be exposed to viruses in the event of inadequate drinking water treatment or due to failures in the distribution network.

Viruses transmitted through water belong to several families, with over 200 types, many of which are associated with epidemics. Families of viruses of priority interest to water belong to enteric viruses. They have the characteristic of being called "naked" viruses, as they consist

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

only of nucleic acid enclosed in a protein capsid that protects it from the external environment, but without lipoprotein outer casing. In recent decades, attention has also been focused on viruses that are mainly responsible for respiratory diseases. These viruses, unlike "naked" viruses, have a pericapsidic envelope (envelope), a structure composed of a double layer of phospholipids and glycoproteins. The two main groups of envelope viruses that may be of concern for the integrated water cycle belong to the families Orthomyxoviridae (influenza virus) and Coronaviridae (SARS and MERS coronavirus). These viruses are known to have been responsible for epidemics and pandemics such as "Spanish" H1N1 influenza (1918-1920), H5N1 avian influenza (1997-today), H1N1 influenza (2009-2010), SARS-cov influenza (2002-2003), MERS-cov (2012), H7N9 avian influenza (2013-today) and, finally, the ongoing pandemic SARS-cov-2 (2020)1. For these groups of viruses, there is currently no evidence of water transmission. However, their presence is demonstrated in the faeces, urine and excreta of infected patients; as a result, viruses can enter the water cycle through wastewater (Figure 39).

However, it is known that, generally, enveloped viruses have much lower persistence characteristics than so-called "naked" viruses, being more susceptible to environmental factors (temperature, sunlight, pH, etc.), as well as physical factors (degree of matrix dehydration) and biological (microbial antagonism). Therefore, even in the absence of specific data on the survival of SARS-cov-2 in water, it is hypothesized that the virus can be deactivated significantly faster than enteric viruses with typical water transmission such as, for example, *Adenovirus, Norovirus, Rotavirus* and *Hepatitis A virus*.

On the basis of the fate of viral pathogens in the integrated water cycle (Figure 40) and the above considerations, three elements of risks analysis during “Pandemic Emergency” are reported in the following portions of the Risk Matrix (Figure 52, Figure 53, Figure 54, Figure 55, Figure 56 and Figure 57).

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

First element of risk analysis: Potential transmission of viruses by wastewater

LINE 2	RISK MATRIX PESCHIERA-BORROMEO WWTP FOR AGRICULTURAL IRRIGATION WATER REUSE (Class B reclaimed water quality)											
		HAZARDS	PRELIMINARY RISK ASSESSMENT									
NODE P-Process T-Transport	WWTP SECTION	HAZARDOUS EVENT (Potential or found events able to introduce one or more specific hazards)	OBSERVATIONS (Based on bibliographic notes, operative personnel, monitoring data, inspections, users complaints)	MICROBIOLOGICAL	CHEMICAL	FISICAL	RADIOLOGICAL	REUSE SERVICE INTERRUPTION	LIKELIHOOD	SEVERITY	SCORE	RISK RATING
ALL NODES	ALL	ELEMENT OF RISK ANALYSIS DURING PANDEMIC EMERGENCY Potential transmission of SARS-CoV-2 virus by wastewater	• The persistence of the virus is negligible and mostly gets destroyed in WWTP ambient temperature and climatic conditions • Only SARS-CoV-2 viral material (RNA) was found in the supply chain • Possible distrust of wastewater safety for spreading of news also not based on evidence	x				x	1	16	16	**High Risk** Hazard or hazardous event potentially resulting in serious illness or injury, or even loss of life and/or will lead to major investigation by regulator with prosecution likely

Figure 52: Preliminary Risk Assessment of potential transmission of SARS-CoV-2 virus by wastewater (first part of the matrix)

RESIDUAL RISK						RE-ASSESSMENT OF THE RISK
EXISTING CONTROL MEASURES	BIBLIOGRAPHIC NOTES - CONSIDERATIONS RELATED TO THE STATISTICAL TREATMENT OF DATA (from D 1.1 + D1.3)	LIKELIHOOD	SEVERITY	SCORE	RISK RATING	INTREGRATIVE CONTROL MEASURES
• P2 Treatments • Equipment maintenance and personal training procedures • Procedures for safety precaution against SARS-CoV-2 exposure for wastewater and sewage workers • Emergency plans • Communication plans connected with the official communication bodies (Guidelines dealing specifically with COVID-19 control in WWTP) • Establishment of a group of experts dedicated to the evolution of knowledge on the subject transmission of viruses through treated water (from WHO website, national institutional websites, Lancet, thematic webinars) in order to adapt control measures and / or update the communication • Establishment of a link between communication experts and the above group of experts		1	1	1	Low risk	• None

Figure 53: Residual risk assessment of potential transmission of SARS-CoV-2 virus by wastewater (second part of the matrix)

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Second element of risk analysis: Dysfunctions related to the unavailability of effective service personnel for management and surveillance

LINE 2	RISK MATRIX PESCHIERA-BORROMEO WWTP FOR AGRICULTURAL IRRIGATION WATER REUSE (Class B reclaimed water quality)											
		HAZARDS	PRELIMINARY RISK ASSESSMENT									
NODE P-Process T-Transport	WWTP SECTION	HAZARDOUS EVENT (Potential or found events able to introduce one or more specific hazards)	OBSERVATIONS (Based on bibliographic notes, operative personnel, monitoring data, inspections, users complaints)	MICROBIOLOGICAL	CHEMICAL	FISICAL	RADIOLOGICAL	REUSE SERVICE INTERRUPTION	LIKELIHOOD	SEVERITY	SCORE	RISK RATING
ALL NODES	ALL	**ELEMENT OF RISK ANALYSIS DURING PANDEMIC EMERGENCY** Dysfunctions related to the unavailability of effective service personnel for management and surveillance (also in the control room in remote control). In particular reduction of monitoring on chemical and microbiological contaminants and on operational variables due to the limitation of human and instrumental resources and external services (e.g. calibration and maintenance of on-line instruments)	• Observed event	x	x	x	x	x	4	8	32	**High Risk** Due to the major regulatory non-compliance

Figure 54: Preliminary Risk Assessment of dysfunctions related to the unavailability of personnel for management and surveillance (first part of the matrix)

RESIDUAL RISK						RE-ASSESSMENT OF THE RISK
EXISTING CONTROL MEASURES	BIBLIOGRAPHIC NOTES - CONSIDERATIONS RELATED TO THE STATISTICAL TREATMENT OF DATA (from D 1.1 e D1.3)	LIKELIHOOD	SEVERITY	SCORE	RISK RATING	INTREGRATIVE CONTROL MEASURES
• Emergency plans (alternative procurement) • SCADA enhancement (eg. personal tablets) • Replacement plans with recruitment in other company sectors • Use of subcontracting for analytical checks • Reduction of staff in service, smart-working incentives, extraordinary standard operating procedure to reduce the possibility of contagion • Definition of indispensable processes, essential services, priority of interventions for the plant • Reinforcement on treatments • Evaluation on historical data • Internal laboratory staff shifts (ensuring the absence of contacts between shift operators) • Agreements with other WWTP managers		1	8	8	**Medium** Due to the major regulatory non-compliance and the reuse service interruption (legal complaints and concern)	• Building of water reservoirs to avoid reuse service interruption • Supply of Class B treated water from other CAP WWTP [see D1.4] • Emergency plans (alternative procurement) • SCADA enhancement (eg. personal tablets) • Replacement plans with recruitment in other company sectors • Use of subcontracting for analytical checks • Reduction of staff in service, smart-working incentives, extraordinary standard operating procedure to reduce the possibility of contagion • Definition of indispensable processes, essential services, priority of interventions for the plant

Figure 55: Residual Risk Assessment of dysfunctions related to the unavailability of personnel for management and surveillance (second part of the matrix)

digital-water.city
digitalwater_eu
digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Third element of risk analysis: dysfunctions related to the unavailability of materials, products and reagents

LINE 2	RISK MATRIX PESCHIERA-BORROMEO WWTP FOR AGRICULTURAL IRRIGATION WATER REUSE (Class B reclaimed water quality)												
		HAZARDS	PRELIMINARY RISK ASSESSMENT										
NODE P-Process T-Transport	WWTP SECTION	HAZARDOUS EVENT (Potential or found events able to introduce one or more specific hazards)	OBSERVATIONS (Based on bibliographic notes, operative personnel, monitoring data, inspections, users complaints)	MICROBIOLOGICAL	CHEMICAL	FISICAL	RADIOLOGICAL	REUSE SERVICE INTERRUPTION	LIKELIHOOD	SEVERITY	SCORE	RISK RATING	
ALL NODES	ALL	ELEMENT OF RISK ANALYSIS DURING PANDEMIC EMERGENCY Dysfunctions related to the unavailability of materials, products and reagents	• Observed event	x	x	x	x	x	4	8	32	High Risk Due to the major regulatory non-compliance	

Figure 56: Preliminary Risk Assessment of dysfunctions related to the unavailability of materials, products and reagents (first part of the matri

RESIDUAL RISK						RE-ASSESSMENT OF THE RISK
EXISTING CONTROL MEASURES	BIBLIOGRAPHIC NOTES - CONSIDERATIONS RELATED TO THE STATISTICAL TREATMENT OF DATA (from D 1.1 e D1.3)	LIKELIHOOD	SEVERITY	SCORE	RISK RATING	INTREGRATIVE CONTROL MEASURES
• To increase as possible the warehouse supplyes • Agreements with other WWTP managers • Requests to civil protection for essential supplies, including personal protective equipment • Expansion of supplier lists (various geographical areas) • Subcontracting		1	8	8	Medium Due to the major regulatory non-compliance and the reuse service interruption (legal complaints and concern)	• Building of water reservoirs to avoid reuse service interruption • Supply of Class B treated water from other CAP WWTP [see D1.4]

Figure 57: Residual Risk Assessment of dysfunctions related to the unavailability of materials, products and reagents (second part of the matrix)

In order to reduce the residual risk score of the last two hazardous events, the following supplementary control and mitigation measures may be applied:

- In the medium to long term, building of water reservoirs to avoid reuse service interruption
- Supply of Class B treated water from other CAP WWTPs [Deliverable D1.4] should be implemented.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Risks prioritization for the SSP at Peschiera Borromeo WWTP

The whole semi-quantitative risk assessment for the reuse system located in the peri-urban district of Peschiera-Borromeo is reported in the Risk Matrix (Annex A).

The output of the semi-quantitative risk assessment, and particularly of the residual risk assessment stage that takes into account the already existing control measures, was a priority-sorted risks list. This list can be very useful to highlight “high” or “very high" risks that need to be quickly faced by the introduction of new specific control measures.

In Figure 58 is reported the numerousness of the hazardous events identified at Peschiera-Borromeo reuse system that belong to different classes of risk (i.e., low risk, medium risk, high risk and very high risk) according to the results of the residual risk assessment.

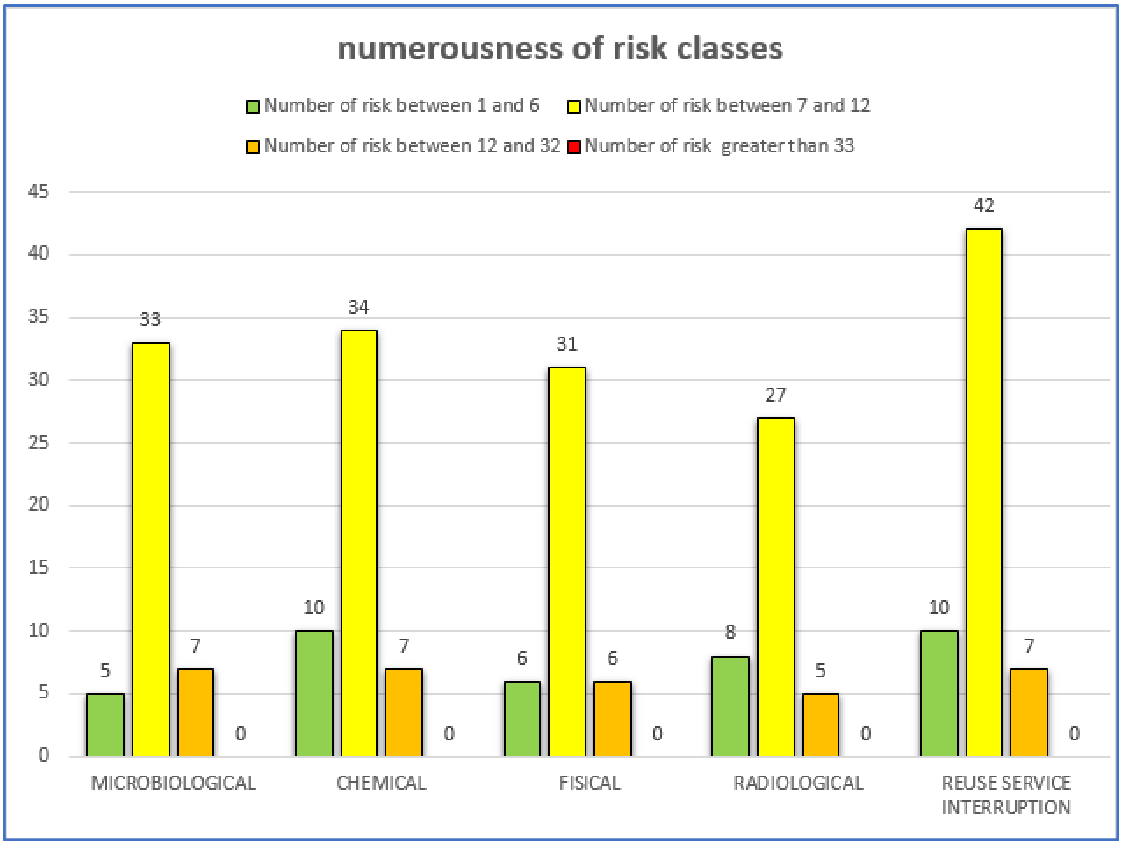


Figure 58: Distribution of residual risk values by different types of hazard

Particularly, the histograms in Figure 58 clearly indicate that most of resultant residual risks are ranked at *Medium Risk* level, no residual risks is ranked as *Very High Risk*, whereas few hazardous events are related to *Low Risk* level (< 6 of risk score R) or *High Risk* level. Hence, to further reduce risks related to the production of water quality of class “B” for agricultural irrigation some integrative control measures can be introduced. One effective additional control measure is represented by the implementation of an Early Warning System (EWS) and

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

its digital infrastructure at Peschiera-Borromeo WWTP. Indeed, a digital EWS tool is able to provide promptly warning messages to assure a safe water reuse.

3.2.5. Quantitative risk assessment

3.2.5.1. QMRA

From the semi-quantitative methodology, which resulted in the Risk Matrix of Peschiera Borromeo wastewater treatment plant (WWTP), a quantitative approach was then followed to analyse microbial risk. A Quantitative Microbial Risk Assessment (QMRA) was carried out to quantitatively evaluate microbial risks and compare them with health targets provided by World Health Organization (WHO) guidelines (WHO, 2016b), which identify four steps:

- Hazard identification: This step involves deciding which microorganisms are of interest in the study and finding out what diseases these microorganisms cause. It includes general information about the microbial agent (pathogens) and the adverse consequences to the host from infection.
- Hazard characterization: The scope of this step is to have a detailed description of the mechanisms and the cause of the actual adverse health effects. It includes information on the required level of detail, hazards (pathogens) and health outcomes to be considered and exposure pathways and hazardous events to be included.
- Exposure assessment: The purpose of exposure assessment is to predict the fate of a hazard from its source to the endpoint and quantity what this endpoint is exposed to. Different groups of people could be exposed to hazards through different pathways. In this step the frequency and magnitude of exposure to pathogens via the pathways and hazardous events are defined. It is important to determine dose-response relations.
- Risk characterization: The final step of the risk assessment combines the information from the previous steps to estimate the likelihood of an adverse consequence (Haas et al., 2014). The health impact data for the identified hazards and the specific study population are reported. They include: the type of health effects, the severity and duration of a disease or illness that may occur after ingestion of the pathogen and the relationship between ingested dose and the probability that health effects (infection, illness, sequelae) occur (WHO, 2016b). Risk characterizations range from a "point estimate" of risk to more sophisticated methods, also known as probabilistic risk assessment, that consider uncertainty in model input parameters and variability across individuals and subpopulations.

Risk assessment includes a certain level of uncertainties within its estimation. The terms variability and uncertainty refer to imprecise or not reliable data, assumptions or lack of information, which might lead to errors in the overall result (Wolfgang Seis, 2012).

Early approaches of quantitative microbial risk assessment (QRMA) were based on point estimates and resulted in a single value of risk. Point estimates means that one value is chosen to represent each variable and the risk is calculated considering that value. Mean values of variables are chosen to

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

calculate the average risk while extreme values, such as the 95-percentile, could give an idea of the worst-case scenario. Such an approach does not give a comprehensive picture nor appropriate weight of all combinations. On the other side, stochastic modelling could be used to have a more realistic representation of the distribution of data. Monte Carlo method can be used to obtain the output risk distribution using random samples of each distribution. This latter approach was used in this study.

In this study, QMRA was performed through the web tool QMRA.org (Wolfgang Seis 2022-QMRA (Version 0.1.3) [Computer software] https://doi.org/10.5281/zenodo.6457511). The tool includes a complete database of typical pathogens concentrations, treatment log removals, exposure scenarios, dose-response and health parameters. Moreover, it can be customised to allow the user to consider site-specific data and/or conditions. The tool QMRA.org runs Monte Carlo simulations and it is able to evaluate the probability of infection and the DALYs associated with defined scenarios of microbiological contamination of reused wastewater.

Hazard identification

The first step of a QMRA procedure is to identify the possible hazardous pathogens that may be reason of illness for humans due to wastewater reuse practices.

Risk assessment is generally conducted for relevant reference pathogens, which should be representative of each the major groups of pathogenic organisms (i.e., bacteria, viruses, protozoa and helminths), since different groups of pathogens may have different behaviours and susceptibilities during wastewater treatment processes.

Reference pathogens were selected considering suggestions by WHO guidelines and by the Australian Guideline for water recycling (NRMMC-EPHC-AHMC, 2006). Hence, *Campylobacter*, *Cryptosporidium* and *Rotavirus* were selected to be representative organisms of bacteria, protozoa and virus, respectively.

Campylobacter infectious are currently the second most frequently reported bacterial infections causing gastroenteritis (Seis et al., 2012). *Campylobacter* bacteria are able to survive in the environment for a certain time, and infections can occur due to the consumption of and bathing in contaminated water. Hence, they can be considered suitable representative of bacteria group to conduct risk analysis.

Rotavirus is a good candidate to be representative of viruses during risk analysis. Indeed, *Rotavirus* poses a major threat of viral gastroenteritis worldwide, has a relatively high infectivity compared with other waterborne viruses and a dose–response model is available in the literature (A.H. Havelaar, J.M.Melse, 2003)

Cryptosporidium parvum is a good candidate to be the reference organism for protozoa. It is reasonably infective, resistant to chlorination, and it is regarded as one of the most significant waterborne human pathogens in developed countries since it is cause of bacterial gastroenteritis (NRMMC-EPHC-AHMC, 2006).

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Hazard characterization

Pathogens concentrations available at Peschiera Borromeo WWTP are related to the concentration of the indicator organisms *E. Coli* in the raw influent to the plant, in the secondary effluent and in the final effluent after UV disinfection, which are measured by periodical standard routine analysis.

To define the expected concentrations of *Campylobacter*, *Cryptosporidium* and *Rotavirus*, typical ratios between *E. coli* and reference pathogens in raw wastewater were used (Mara, 2006), as reported in Table 21.

Table 21 Reference ratios between *E. Coli* and pathogens concentration in raw wastewater

***E. coli* - reference ratio**			
pathogen	***Campylobacter***	***Cryptosporidium***	***Rotavirus***
min	1.00E-06	1.00E-07	1.00E-06
max	1.00E-05	1.00E-06	1.00E-05

Particularly, for Cryptosporidium and Rotavirus the literature ratios shown in Table 21 were applied to the concentration of *E. coli* in the influent. Literature data on log removals at each treatment unit were then applied to define the pathogens concentrations in the effluent.

Reference values (NRMMC-EPHC-AHMC, 2006; Wolfgang Seis & Remy, 2013; WHO, 2006) for the applied log reductions are reported in Table 22.

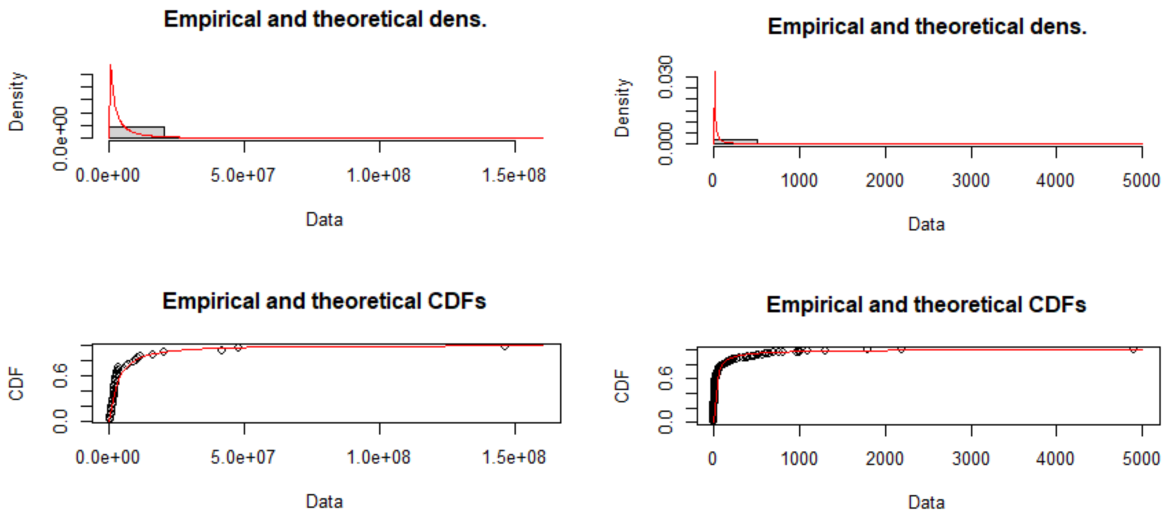
Table 22: Indicative log-removals for *Cryptosporidium* and *Rotavirus* in different wastewater treatment units

Treatment unit	*Distribution*	*Cryptosporidium*		*Rotavirus*	
		min	max	min	max
Primary Treatment	uniform	0	1	0	0.1
Secondary Treatment	uniform	0.5	1.5	0.5	2
UV disinfection	uniform	3	3	1	3

In the case of *Campylobacter*, the literature ratios shown in Table 21 were applied to the concentrations of *E. Coli* in the final effluent of Peschiera Borromeo WWTP. Indeed, *Campylobacter* and *E. Coli* are both bacteria and, thus, same log removals are applicable in each treatment unit for the two organisms (NRMMC-EPHC-AHMC, 2006; Wolfgang Seis & Remy, 2013; WHO, 2006).

E. Coli concentrations in both the influent and effluent of the WWTP were fitted by a log-normal distribution (Figure 59). Then, uniform distributions were utilized to apply the ratios in Table 21 and obtain the needed distribution of *Cryptosporidium*, *Rotavirus* and *Campylobacter*.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.



a) b)

Figure 59: Fitting of *E. coli* trends in the a) influent, and b) effluent with lognormal distributions

Exposure Assessment

In this step, health impact data for the identified hazards and the target population are specified. This includes the type of health effects, the severity and duration of a disease or illness that might occur after ingestion of the pathogen, and available information on the relationship between ingested dose and the probability that health effects occur (dose–response relationship). In addition, the fraction and vulnerability of the population exposed might need to be considered.

Fieldworkers and local communities might be exposed through direct contact with wastewater or contaminated soil or crops. These groups might inhale or ingest wastewater used for irrigation. Each group could enter in contact with contaminated wastewater or products in several ways or routes of exposure. Different barriers or log removals could be applied at each route of exposure to reduce the risk.

Other possible exposure routes of humans to pathogens due to wastewater irrigation could be considered, but not all of them are relevant for the case-study in object. As an example, the exposure via drinking water can be neglected, since there are no contacts with water catchment areas for providing drinking water.

Moreover, in Peschiera Borromeo reuse system, no crops are grown to be consumed without further processing. Thus, the pathways through the consumption of animals and animal products as well as direct consumption of crops grown on wastewater irrigated areas are neglected, too.

In conclusion, both fieldworkers and local community were considered as exposure groups through ingestion and inhalation. No pathogen die-off was assumed after irrigation. Dose-

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

response models proposed by Haas were adopted to evaluate the risk of infection per person per year (pppy).

Routes of exposure

The routes of exposures are the activities through which people enter in contact with pathogens.

A volume (mL) and a frequency per person per year (pppy) of exposure must be associated to each route. Due to lack of data and impossibility to obtain site-specific information, suggestions provided by Australian Guidelines were considered. The main route of exposure to microbial hazards from recycled wastewater is ingestion, including ingestion of droplets produced by sprays (inhalation). Dermal exposure is also possible, but there is a lack of evidence of health impacts through this route and it is considered unlikely to cause significant levels of infection or illness in the normal population (NRMMC-EPHC-AHMC, 2006). Data on exposure volumes and frequencies per person, utilized in this study, are provided in Table 23 according to the indications of the Australian water recycling Guideline (NRMMC-EPHC-AHMC, 2006) and Seis et al. (2012).

Table 23 Associated exposures for recycled water during irrigation (NRMMC-EPHC-AHMC, 2006)

Activity	Exposed group	Rout of exposure	Volume (mL)	Frequency/person/year
Agriculture irrigation	Fieldworker	Indirect ingestion	0.01	100
	Nearby community	Inhalation (min)	0.0045	180
		Inhalation (max)	0.0069	180
	Children in the nearby community	Ingestion (min)	0.02	10
		Ingestion (max)	0.1	10

For highly mechanized agriculture, a daily intake of 1-10 mg contaminated soil, or 1-10 µl treated wastewater can be assumed for fieldworkers with a number of exposure events per year of 100 days per person per year (Seis et al., 2012). In this case, it is assumed that number of 100 ml water ≙ number of 100 g soil (Seis et al., 2012).

For assessment of the exposure of nearby residents, the dose of liquid aerosol particles people are exposed to must be estimated. Due to lack of local data, it can be considered that, due to wind presence, nearby resident could inhale 4.5 - $6.9*10^{-3}$ mL of water, in the case of spray irrigation, for each irrigation event (Seis et al., 2021). The number of exposure events per year is equal to the number of irrigation events in one year. In the peri-urban area of Milan, agricultural fields cultivated with corns are irrigated once per day for six months. Hence, an exposure of 180 events per person per year can be assumed. However, spray irrigation is not applied in the

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

investigated area and this exposure scenario does not result in a high risk for the local community.

On the contrary, children of the nearby community may play on agricultural areas or may accompany adults while they go for a walk. Especially young children tend to ingest higher amounts of soil. To account for this kind of risk an annual number of exposure events of 10 is applied. The amount of soil ingested is set to 20-100 mg (i.e., 0.02 – 0.1 mL of water) per exposure event according to Seis et al. (2012).

Reassuming, in this study the (indirect) ingestion of wastewater by fieldworkers and by children of the local community were considered as the most relevant exposure routes for microbial risk calculation.

Preventive measures

Preventive measures, or barriers, are strategies to reduce the exposure to hazard, which include actions to reduce the pathogens concentration in the wastewater (e.g., wastewater treatment processes) or actions to reduce the volume of water target people are exposed to.

Each preventive measure can be associated with a log reduction value. Characteristic values of log reduction have been taken from (NRMMC-EPHC-AHMC, 2006) and WHO Guidelines (2006) and are summarized in Table 24.

Table 24 Log reductions applied to each barrier

Reference	*NRMMC-EPHC-AHMC, 2006*		WHO guidelines for safe use of wastewater, excreta and greywater	
	min	max	*min*	max
Drip irrigation of crops	2			
Drip irrigation of crops with limited to no ground contact (e.g., tomatoes, capsicums)	3			
Drip irrigation of raised crops with no ground contact (e.g., apples, apricots, grapes)	5			
Drip irrigation of plants/shrubs	4			
No public access during irrigation	2			
No public access during irrigation and limited contact after (until dry 1 – 4 hours) (e.g., food crop irrigation)	3			
Buffer zones (25–30 m)	1		1	
Natural die-off			0.5	2

In the present study, a log removal value of 2 was considered to evaluate the exposure when drip irrigation is utilized. Indeed, the use of drip irrigation significantly reduces the volume of utilized wastewater and the production of aerosol compared to other irrigation technologies (i.e., spray irrigation or surface irrigation).

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

For the selected exposed groups, two microbial risk scenarios were analysed, which were related to the application or not of drip irrigation.

Dose-response relationship

The dose of pathogens with which the exposed group enter in contact can be calculated with the following equation:

$$d = \frac{c * exposure/event}{\log reduction} \qquad (3.1)$$

Where:

- C = concentration of pathogens
- log reduction = reductions required to achieve a residual risk coming from preventive measures or barriers
- exposure/event= volume (mL) with which people enter in contact in a single event of exposure during a certain activity

Next step consists in the identification of dose-response models, which are mathematical functional relationships between the number of pathogens someone is exposed to and the probability of occurrence of the related adverse effect. A fraction of infected people may develop different health outcomes. The simplest dose-response model is an exponential relationship (Haas et al., 2014):

$$P_{inf} = 1 - e^{-r*d} \qquad (3.2)$$

- P_{inf} =probability of infection
- r= infectivity constant
- d= dose

The exponential model assumes that the probability of infection is constant for similar kind of pathogens (C.N. Haas et al, 1999), while actually not all pathogens of the same species are equally infective and not all humans have the same health outcomes. For that reason, the Beta Poisson-model is also used (Haas et al., 2014):

$$P_{inf} = 1 - (1 - \frac{d}{\beta})^{-\alpha}$$

(3.3)

Where:

- *d = dose per event*
- *α, β = dose response constants*

A Dose-response model is used to calculate the Probability of infection (P_{inf}) related to each event of exposure to the target pathogen. In this study, according to what suggested by (Haas et al., 2014), a Beta-Poisson model has been used to calculate P_{inf} in the case of *Campylobacter* and *Rotavirus*, while an exponential model was used for *Cryptosporidium*.

$P_{inf} = 1 - e^{-r*d}$ (used for *Campylobacter* and *Rotavirus*)

$P_{inf} = 1 - (1 - \frac{d}{\beta})^{-\alpha}$ (used for *Cryptosporidium*)

Constant values "α" and "β" for the Beta-Poisson model and "r" for the Exponential equation were obtained from literature (Table 25). Particularly, reference values were obtained by Australian guidelines, QMRA Wiki community portal (a data repository website created by researchers of the Center for Advancing Microbial Risk Assessment (CAMRA) OF Michigan State University), and Aquanes tool website.

Table 25 Dose-response constants for selected pathogens

		Campylobacter	***Cryptosporidium***	***Rotavirus***
Dose-response constants	**α**	1.44E-01		2.53E-01
	β	7.58		4.26E-01
	r		5.72E-02	

Risk characterization

Once calculated Pinf, the total probability of infection in one year is obtained by the following equation:

$$\mathrm{P_{inf}\ combined\ final} = 1 - \prod_1^n (1 - \mathrm{P_{inf\,i}})^{\mathrm{Ni}} \quad (3.7)$$

Where:

- N = number of activities
- Pinf i = probability of infection of the i^{th} activity
- Ni = frequency/person/year of i^{th} activity

The final probability of infection of each pathogen calculated by eq. (3.7) should be compared with the limit value established by US EPA of $1*10^{-4}$. If the calculated probability of infection is lower than this health-target value, the microbial risk can be considered acceptable.

Once the value "$\mathrm{P_{inf}}$ combined final" has been calculated, the probability of illness (P_{ill}) is obtained multiplying "$\mathrm{P_{inf}}$ combined final" with the ratio illness/infection, which is provided by literature (Table 26) (NRMMC-EPHC-AHMC, 2006):

$$\mathrm{Pill} = \mathrm{P_{inf}\ combined\ final} * \mathrm{ratio\ illness/infection} \quad (3.8)$$

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Table 26 Ratios illness/infection for the selected pathogens

Campylobacter	*Cryptosporidium*	*Rotavirus*
0.3	0.7	0.5

WHO guidelines consider Disability-adjusted life years (DALYs), as a metric for expressing the burden of disease within a population. The DALY is a health gap indicator for the status of health of a population expressed as burden of disease due to a specific disease or risk factor, and it takes into account both the morbidity and the mortality caused by a specific disease. A health-target of 10^{-6} DALYs was set by WHO as tolerable health risk.

The DALY value can be calculated by the following equation:

$$\text{DALY per year} = \text{Pill} * \text{DALYd} * \text{susceptibily fraction} \quad (3.9)$$

Where:

- Pill= probability of illness per year
- DALYd= DALY per case

Daly per case of reference pathogens were obtained by WHO guidelines (2011), as shown in Table 27.

Table 27: DALY per case for the selected pathogens

	Campylobacter	***Cryptosporidium***	***Rotavirus***
DALY per case	0.0046	0.0015	0.014

Once values of DALYs per year for each pathogen have been calculated, the obtained values can be compared with the tolerable level of risk ($1*10^{-6}$) set by WHO in order to understand if the risk is acceptable or not.

Results

Results are summarised in the following paragraphs, both for workers and for children of the local community.

Graphs (Figure 60 – Figure 65) show the calculated Probability of infection that were compared with the health-target limit suggested by US EPA (10^{-4}) and the calculated DALYs. In this case, the tolerable value of 10^{-6} suggested by WHO Guidelines was considered as the threshold value.

In each graph, results of the expected microbial risk are shown considering minimum and maximum log removals that can be achieved during wastewater treatment. Indeed, log removals for viruses and protozoa during the different treatment stages were taken from literature and are highly variable since are referred to WWTPs that work under very different operational conditions (Table 22). Particularly, low log removals are often related to small WWTPs that work under non-optimised conditions and that reach low treatment performances. On the contrary, high values of log removals comes from large WWTPs, which are generally operated under better conditions. Showing results for minimum and maximum removal of pathogens gives indication of the uncertainty of the assumptions done to evaluate

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

the risk. In these terms, the variability of the obtained results (i.e., difference between minimum and maximum log removal scenarios) represents the uncertainty related to the use of literature data and assumptions done to evaluate the risk. Only site-specific measurement campaigns could confirm or deny the assumed hypothesis, allowing to utilize site specific values for expected log removals of pathogens at Peschiera Borromeo WWTP.

A decision support tools such as the Early Warning System will be then very useful to lower the expected microbial risk. Indeed, the EWS can detect plant malfunctions or low treatment efficiencies and rapidly send alarms. Hence, the presence of the EWS can ensure that when wastewater is provided for irrigation, the WWTP is working properly and that treatment performances met the required levels. In this case, risk scenarios related to high log removals of pathogens will result more plausible.

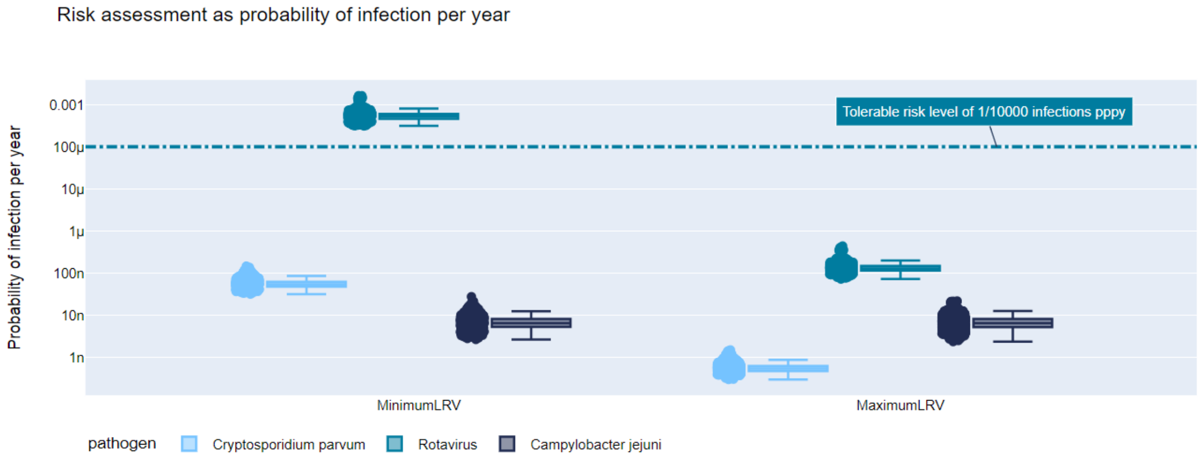
Risk for workers

Risk was evaluated for fieldworkers, considering or not drip irrigation as a barrier. Indeed, wastewater treated at Peschiera Borromeo WWTP is reused in agriculture for corn cultivation by surface irrigation and by drip irrigation.

Figure 60 and Figure 61 show the calculated microbial risk in the case of use of drip irrigation and conventional irrigation, respectively. Results are similar for both the scenarios. Particularly, the probability of infection and DALY values are higher than related health-target limits only for the pathogen *Rotavirus* in the case of low performance of wastewater treatment operations and low log removals for pathogens. Hence, the uncertainty related to literature values for log removals (Table 22) does not ensure that an acceptable level of risk is reached during wastewater irrigation.

A strategy to reduce uncertainties about treatment performances could be represented by the implementation of an Early Warning System. If the tool is operated, it can provide information about the status of the wastewater treatment units and thus guarantee that they are working with high treatment efficiency. The presence of the EWS can ensure the good operation of treatment units and that the required performance levels are met, allowing to calculate microbial risk using high log removal for pathogens (Figure 60 and Figure 61).

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.



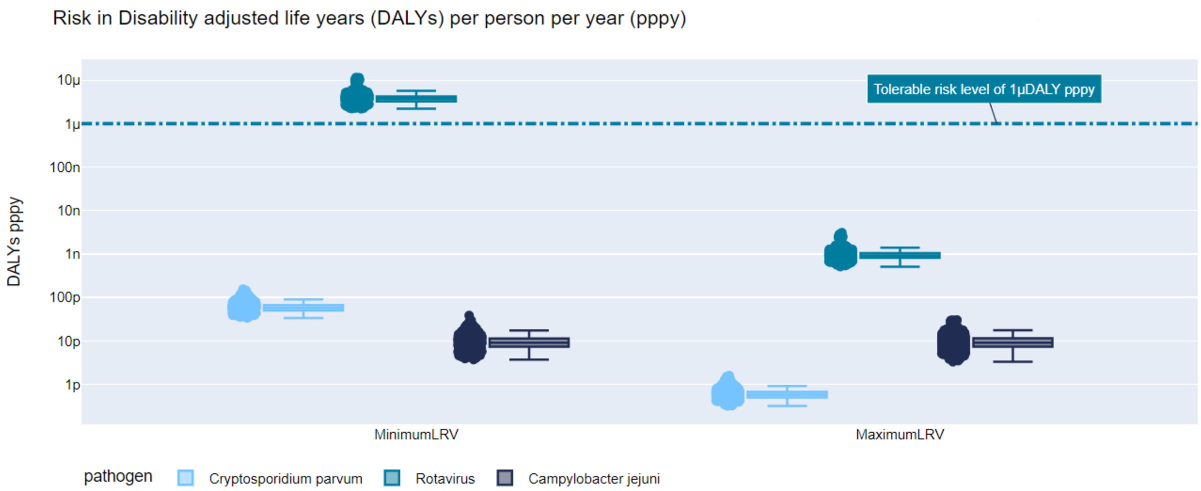


Figure 60: Risk expressed in Probability of Infection (up) and DALYS (down) for workers with drip irrigation in the case of maximum log removals (MaximumLMV) and minimum log removals (MinimumLMV) of pathogens.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

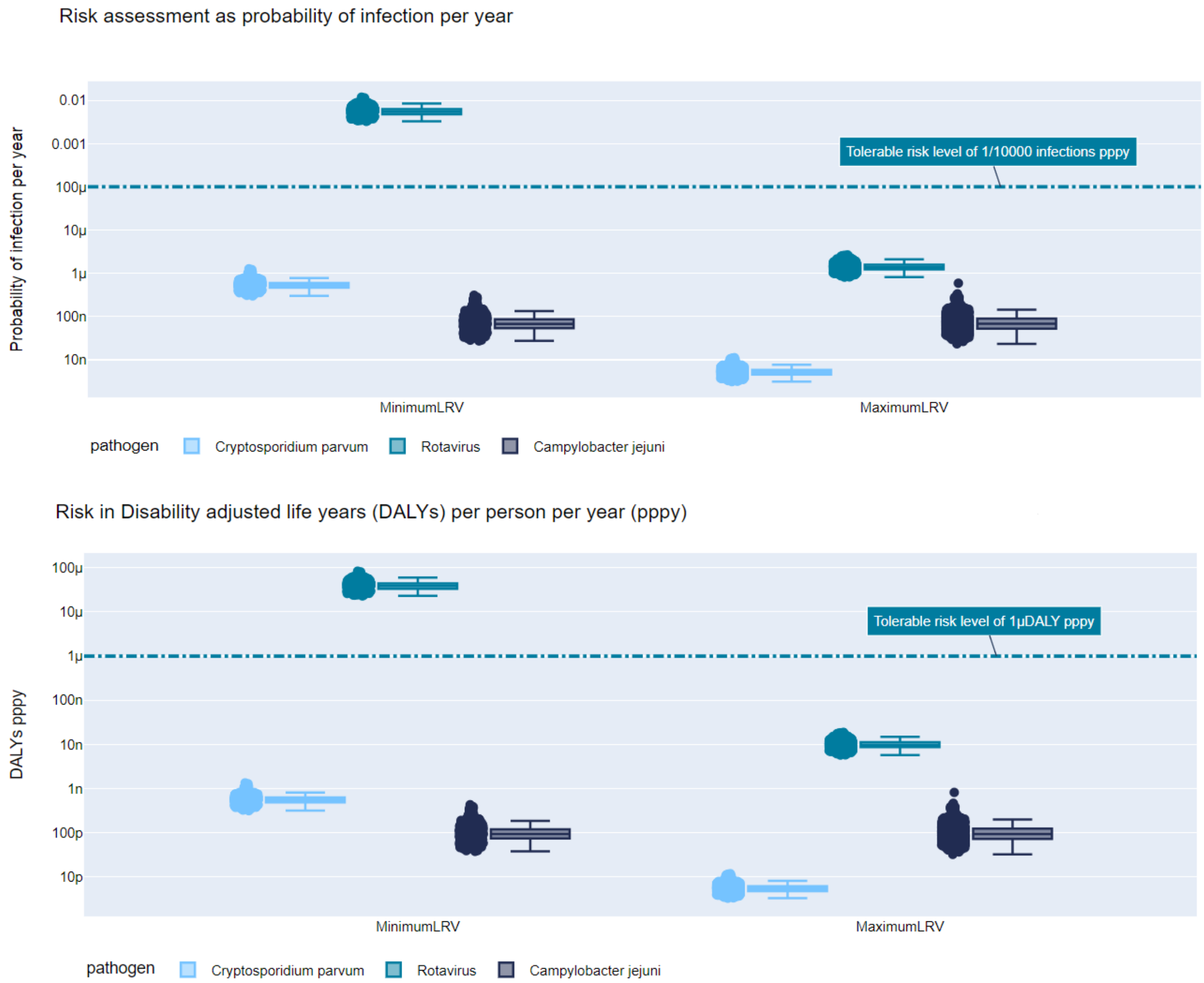


Figure 61: Risk expressed in Probability of Infection (up) and DALYS (down) for workers without drip irrigation in the case of maximum log removals (MaximumLMV) and minimum log removals (MinimumLMV) of pathogens.

Even though for *Rotavirus* the use of drip irrigation was not able to assure an acceptable level of risk for wastewater reuse, this barrier results in any case highly effective to reduce the risk for fieldworkers as it was better highlighted in Figure 31.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Risk assessment as probability of infection per year

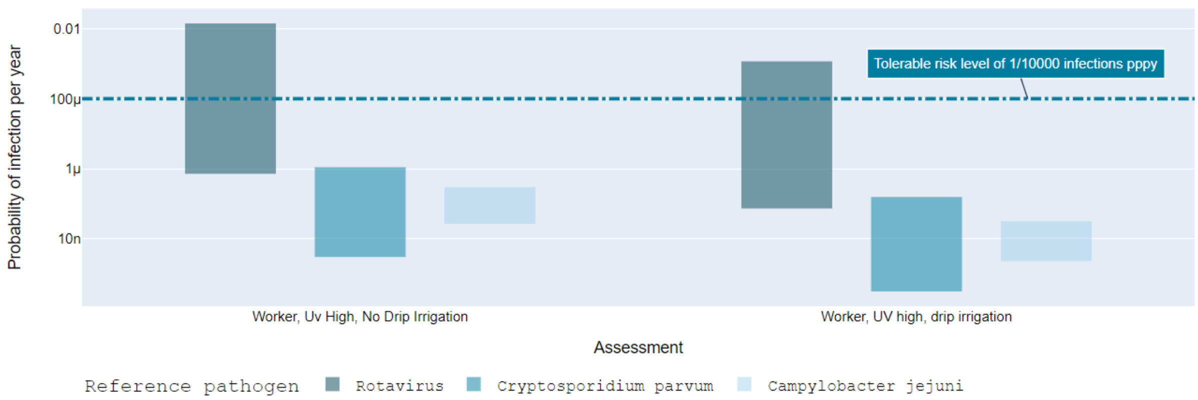


Figure 62: Comparison of risk calculation for workers related to scenarios without and with the use of drip irrigation as a barrier

Risk for local community

Outcomes of risk calculations for children living in the nearby community are reported in Figure 63 and Figure 64. However, in this case, when drip irrigation is utilized for irrigation, the risk levels is always below the health-target threshold limits (Figure 63). On the contrary, if drip irrigation is not used, the risk is not acceptable for *Rotavirus* in the case of low performance of the wastewater treatment processes (MinimumLRV in Figure 64). This latter result highlights again the importance of an EWS able to monitor continuously the quality of the reused wastewater to assure a safe irrigation reuse.

Risk for local community is performed using the same hypothesis as the ones expressed for fieldworkers but considering a lower number of exposure events.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

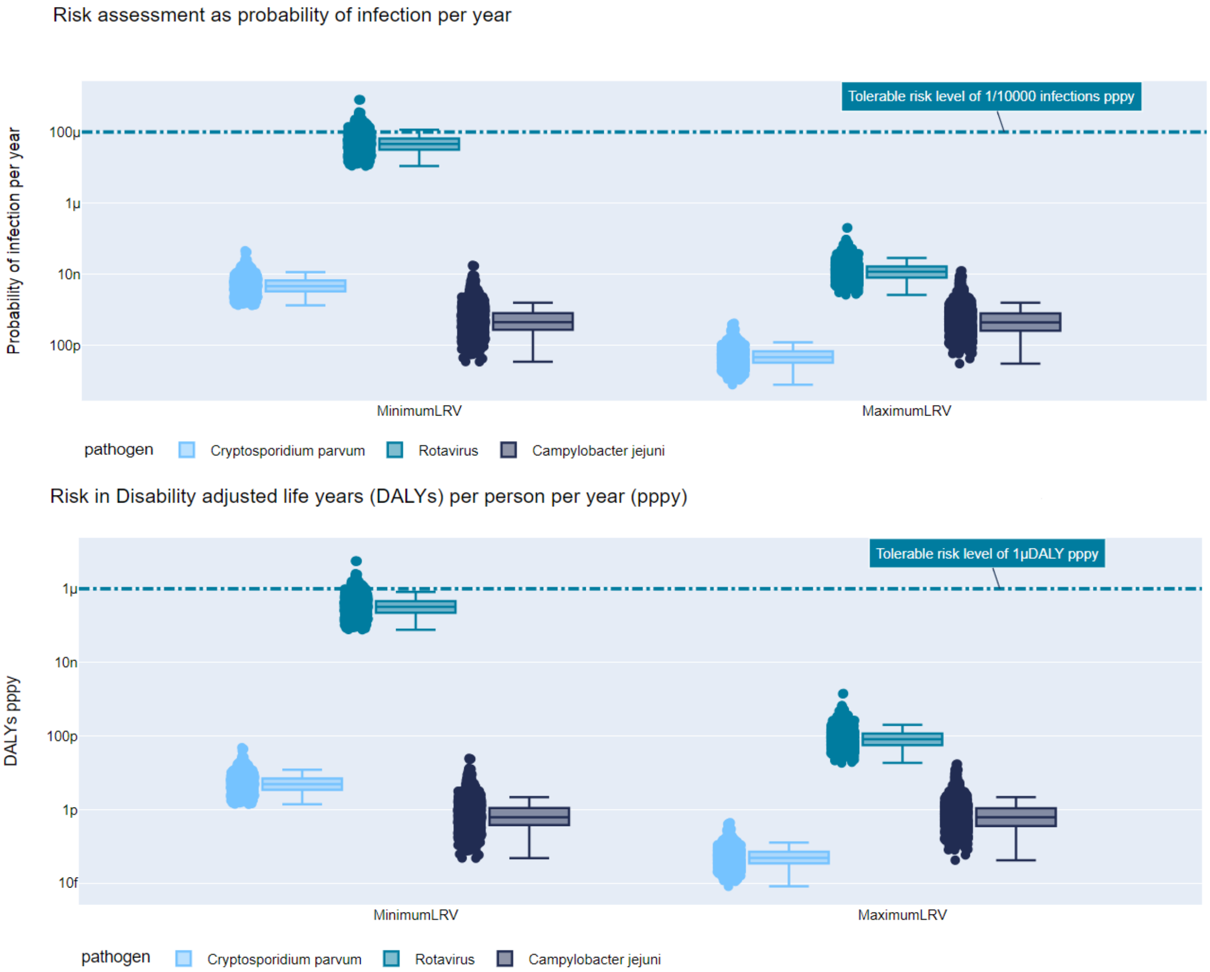


Figure 63: Risk expressed in Probability of Infection (up) and DALYS (down) for local community with drip irrigation in the case of maximum log removals (MaximumLMV) and minimum log removals (MinimumLMV) of pathogens.

digital-water.city
digitalwater_eu
digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

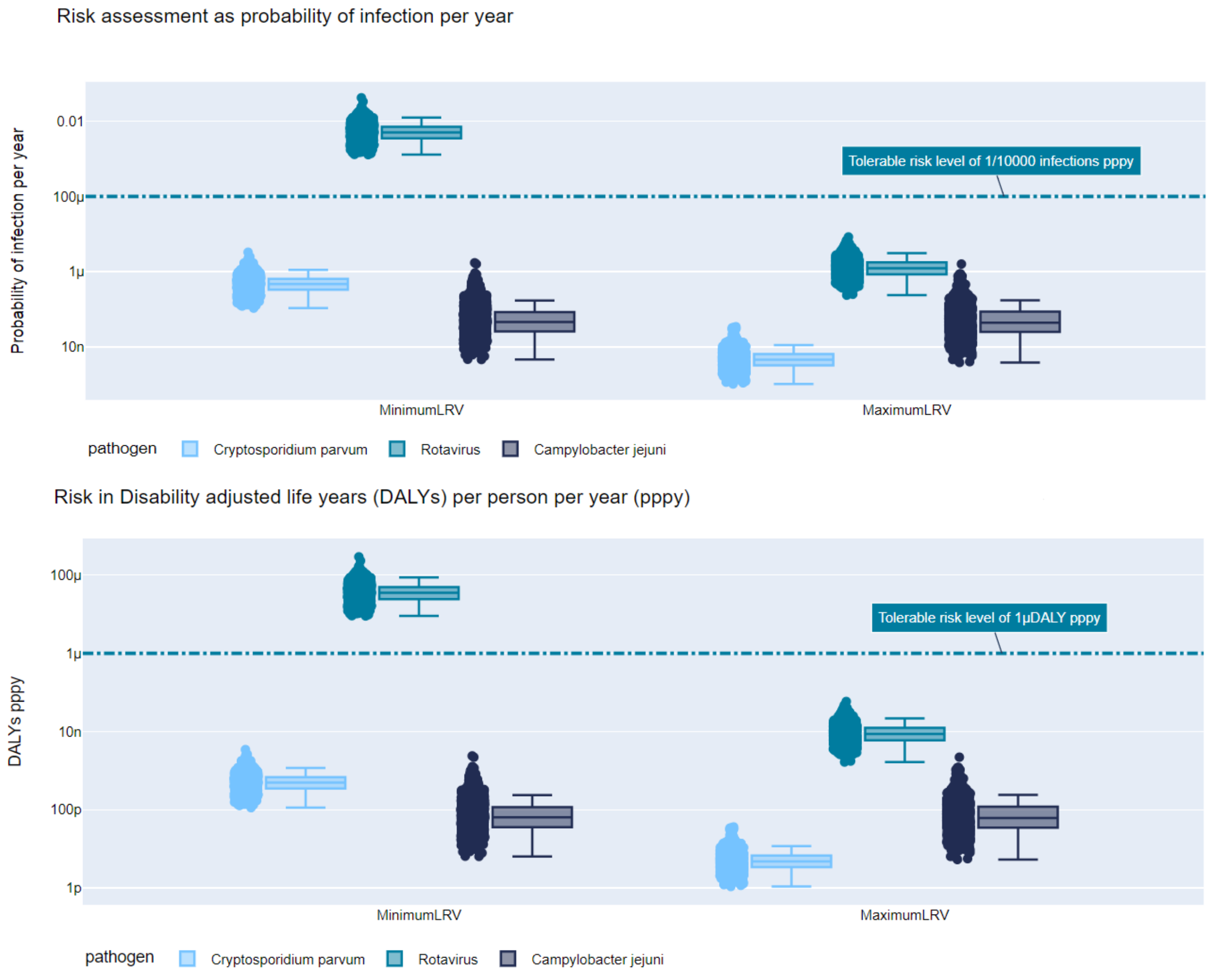


Figure 64: Risk expressed in Probability of Infection (up) and DALYS (down) for local community without drip irrigation in the case of maximum log removals (MaximumLMV) and minimum log removals (MinimumLMV) of pathogens.

The efficiency of drip irrigation as a barrier to reduce risk for local community is shown in Figure 34.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

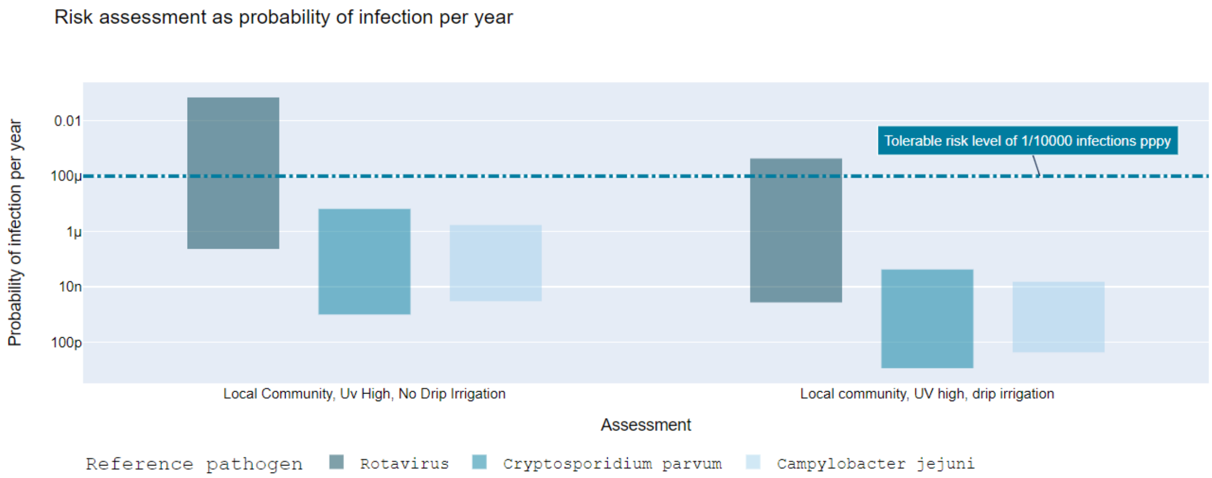


Figure 65: Comparison of risk calculation for local community related to scenarios without and with the use of drip irrigation as a barrier

1.1.3.1. QCRA

Quantitative chemical risk assessment is a tool increasingly used in risk-management decision-making, following the success of its microbiological equivalent (QMRA). In QCRA, available data and information regarding toxicity are combined with estimates of exposure to calculate the likelihood and severity of human health effects. In certain circumstances, limitations on evaluating chemical toxicity and exposure potential introduce significant uncertainties into such risk assessment (Bahri et al., 2010).

According to Annex II of the Regulation on key elements of risk management, the quantitative risk assessment shall be used where there is sufficient supporting data or in projects presenting a high potential risk to the environment or public health. To finalise QCRA of Peschiera –Borromeo case study, a sufficient number of laboratory data, collected by CAP during the monitoring programs reported in paragraph *Sampling and periodical laboratory monitoring*, was used.

The quantitative chemical risk assessment (QCRA) follows the methods of the European Union Technical Guidance Document on Risk assessment (EU 2003). Like QMRA, the QCRA is structured in:

- Hazard identification
- Hazard characterization
- Exposure assessment
- Risk characterization

As for QMRA procedure, the first step was to identify all the possible chemical/physical hazards that could be reason of illness for humans due to wastewater reuse practices. For this purpose, FMEA model was the tool selected.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Hazards identification by FMEA model

FMEA (*Failure Modes-Effects Analysis*) arose, in the industrial field, as a methodology used to identify and evaluate the risk of a potential adverse event in a production process. Therefore, it is a control tool since it studies the hazardous events that may occur, the consequent effects and identifies strategies to reduce the probability that a certain adverse event may reoccur.

This tool can be applied to the study and monitoring of any production chain in order to effectively identify deviations in the quality of goods produced, before they manifest real defects. In our case the good produced is water destined for reuse in agriculture. The production chain is the sequence of management processes that take place from the sewer system, through the treatments to finish at the point of delivery/compliance. Starting from experimental evidence (laboratory analysis) representative of the entire production process of water destined for reuse in agriculture, FMEA analysis allows to identify the critical factors/chemical-physical parameters that have an impact on the quality of the wastewater of sewer system, treated and distributed at the point of delivery. These parameters can lead to different events/phenomena:

- abnormal pollution of the wastewater treatment plant to be detected quickly and managed with appropriate control measures;
- incorrect sampling procedure;
- selection of unsuitable non-representative sampling points (stagnant water sampling)

Finally, this study allows to orient the analytical control plan in the selection of the parameters to be monitored to keep under control the system as well as the most appropriate monitoring frequencies. It is an intuitive graphic model, which associates to each chemical/physical parameter an index "FMEA" (between 1 and 5) that expresses the deviation of its value from the legal limit. In this sense, it offers a snapshot of the quality level of the wastewater treatment and allows to trace the level of chemical risk to which the system is subjected by monitoring its evolution over time.

As part of the Sanitation Safety Plans, the FMEA model allows to:

- identify chemical/physical hazards
- continuously verify the effectiveness of the SSP.

Implementation of FMEA model for Peschiera-Borromeo water reuse system

The implementation of the FMEA model for the production chain of Peschiera-Borromeo treated wastewater for reuse in agriculture was structured according to the following 4 steps.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

- FIRST STEP: selection of the most representative parameters of water quality for reuse in agriculture

The most representative parameters of water quality for reuse in agriculture were chosen by CAP in its monitoring program (paragraph *Sampling and periodical laboratory monitoring*) on the basis of the requirements of the current Italian legislation, Decree 12/6/2003, n. 185 of the Ministry of the Environment on regulation of technical standards for the reuse of wastewater. The chemical and physico-chemical parameters selected by CAP included conductivity, BOD_5, COD, Total Suspended Solids, Total Nitrogen, Ammonium, Nitrate, Total Phosphorous, Phosphates, Chlorides, Sulphates, Sulphites, Sulphides, Cyanides, several heavy metals (Cadmium, Chromium, Iron, Mercury, Nichel, Lead, Copper, Zinc) and organic surfactants. Since for most parameters, data had values below LOQ, FMEA approach was applied to only 13 parameters, which are reported in Table 28 and that are characterized by a consistent number of detections with values above LOQ and have a demonstrated dangerousness. In the future, new parameters will gradually be added to these, depending on scientific evidence and according to new legislations. An example could be represented by the emerging contaminants studied in section *Preliminary studies*.

- SECOND STEP: Calculation of the 95th percentile of the observed values for the selected parameter

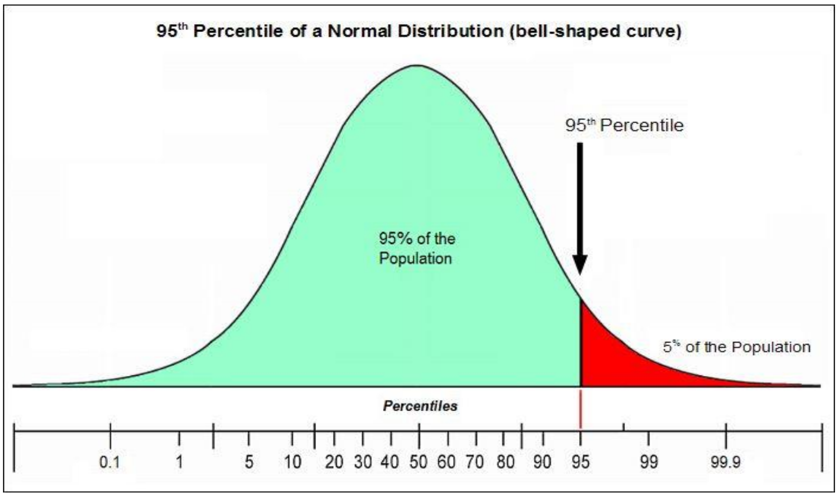


Figure 66: Graphical representation of 95th percentile

To define the FMEA indexes, the 95°percentile has been calculated for all the selected parameters that were measured at concentration higher than LOQ.

The *95th percentile* is a number/value that is greater than *95%* of the numbers/values in a given set of data. The 95° percentile is used to evaluate if the data rank in the range defined by the LL and the LOQ as explained below.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

- THIRD STEP: Definition of the calculation range of the FMEA index

To define the calculation range of the FMEA index, three different limits were considered. For the lower end was considered the LOQ (the minimum quantifiable concentration capable of returning a signal 3 - 2 times higher than the background noise) of each chosen parameter, according to the method used by CAP laboratory for its determination. For the upper end, two different limits for each parameter were considered: one provided by the ISO 16075-1:2020 guideline, and one provided by the Italian DM185/2003 (LL). Generally, limits set by the ISO 16075-1:2020 are less strict than those of the DM185/2003. In the case limits were not available by the ISO 16075-1:2020, limits from *Quaderno ARSIA 5-2004* (ARSIA-L) were considered (Landi & Baroncelli, 2000).
Edge values of the ranges set for the calculation of the FMEA index are reported in Table 28 for all selected parameters.

Table 28: LOQ, LL, ISO-L and ARSIA-L of each parameter

N°	Parameter	LOQ	DM185/2003 limits of Italian law (LL)	ISO 16075-1 limits (ISO-L)	Quaderno ARSIA 5-2004 limits (ARSIA-L)
1	Conductivity at 25°C	150 mS/cm	3000 µS/cm		1500 µS/cm
2	BOD5	5 mg/L	20 mg/L	-	-
3	COD	15 mg/L	100 mg/L	-	-
4	TSS	5 mg/L	10 mg/L		30 mg/L
5	Total Nitrogen	0.1 mg/L	15 mg/L	35 mg/L	
6	Ammonium nitrogen (NH_4)	0.44 mg/L	2 mg/L	30 mg/L	
7	Total phosphorus (P)	0.313 mg/L	2 mg/L	7 mg/L	
8	Chloride (Cl^-)	3.478 mg/L	250 mg/L		200 mg/L
9	Sulfate (SO_4)	2.292 mg/L	500 mg/L		300 mg/L
10	Aluminium (Al)	0.112 mg/L	1 mg/L	12.5 mg/L	
11	Iron (Fe)	0.041 mg/L	2 mg/L		3 mg/L
12	Copper (Cu)	0.010 mg/L	1 mg/L		1 mg/L
13	Zinc (Zn)	0.055 mg/L	0.5 mg/L		3 mg/L

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

FOURTH STEP: Definition of the five classes of the FMEA index and classification of 95 percentile (P95) in each of the classes

For each of the 13 parameters in Table 28, the calculation of the FMEA index was performed according to the criteria reported in Table 29, where formulas are based on the calculation of the 95th percentile (P95).

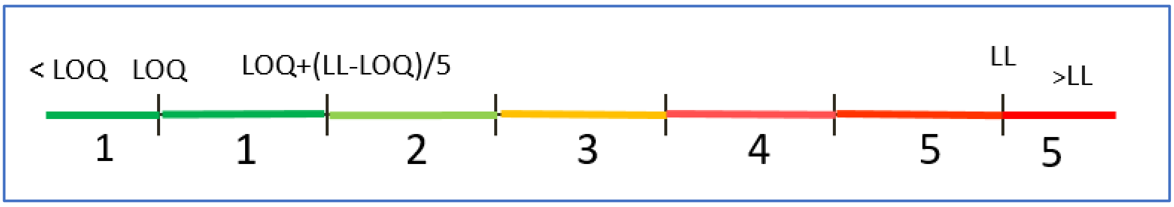


Figure 67 Example of the range of the FMEA index [LOQ, LL] divided into 5 parts numbered from 1 to 5 according to the criteria indicated in Table 29 and characterized by different colours

Table 29: P95 calculation formulae for the five classes of FMEA indices, taking LL as an example of an upper limit

FMEA index	Calculation Formula
I.FMEA = 1	P95 < LOQ + (LL-LOQ)/5
I.FMEA = 2	LOQ+(LL-LOQ)/5 ≤ P95 < LOQ+(2/5)x (LL-LOQ)
I.FMEA = 3	LOQ+(2/5)x(LL-LOQ) ≤ P95 < LOQ+(3/5)x (LL-LOQ)
I.FMEA = 4	LOQ+(3/5)x (LL-LOQ) ≤ P95 < LOQ+(4/5)x (LL-LOQ)
I.FMEA = 5	LOQ+(4/5)x (LL-LOQ) ≤ P95 < LL
I.FMEA = 5	P95 > LL

For all those parameters that had a P95 value included in the first three classes (P95 < 63 % of the legal limit) defined in Table 29 the associated risk can be considered negligible (low or very low). On the contrary, the parameters falling in classes 4 and 5 (P95> 63 % of the limit) need the planning of a careful monitoring program.

Results of FMEA and PCA for QCRA implementation

In order to identify the chemical and chemical-physical hazards within the Peschiera-Borromeo reuse water production chain (Line 2) for QCRA implementation, FMEA and PCA models were applied to lab data collected by CAP.
To implement the FMEA model, the calculation range of the FMEA index was defined considering both the limits provided by the ISO 16075-1:2020 limit (ISO-L) and the *Quaderno ARSIA 5*-2004 (ARSIA-L), and by the Italian DM185/2003, which is the national regulation for

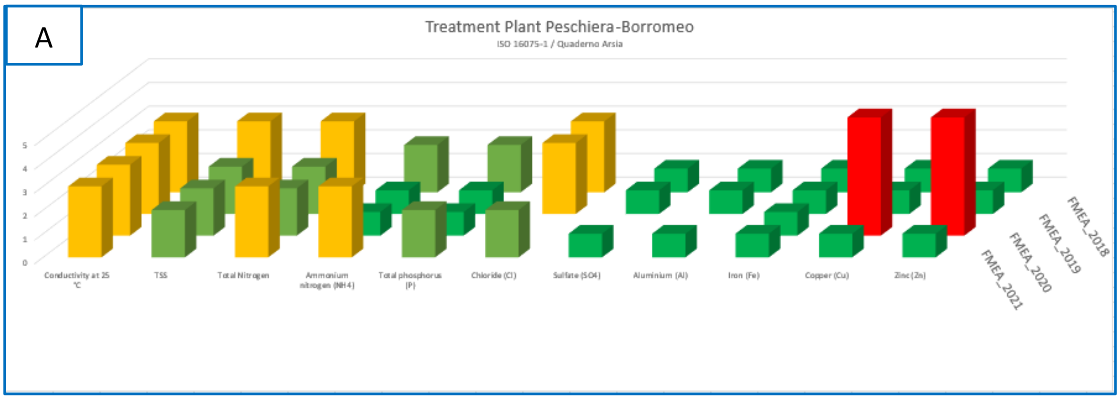
digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

water reuse, as explained previously. Hence, the FMEA index was calculated twice for each parameter to take into account the limits set by both the regulation.

Results of the FMEA model using limits set by ISO-L and ARSIA-L

The histograms in Figure 68 A and B show the FMEA model applied to the data set of analyses carried out in the effluent of Peschiera-Borromeo WWTP (Line 2) between 2018 and 2021. The FMEA index was calculated for each year of the period 2018 – 2020 and for all the selected parameter (Figure 68). In the 3D histogram of Figure 68 A, the level of risk (FMEA index in z-axis) is diversified according to the different parameters (x-axis) and to the different years (y-axis). On the contrary, in the cumulative graph of Figure 68 B, it is possible to observe the percentage of variation of the FMEA index over time. Results from both elaborations highlight the parameters that have to be considered as priority for monitoring/assessment purposes.

In 2021, all investigated parameters fall into classes 1, 2 and 3, where the associated risk can be considered negligible (low or very low). The same results were found in the previous three years, with the exception of copper and zinc, which had a FMEA index 5 for 2020.



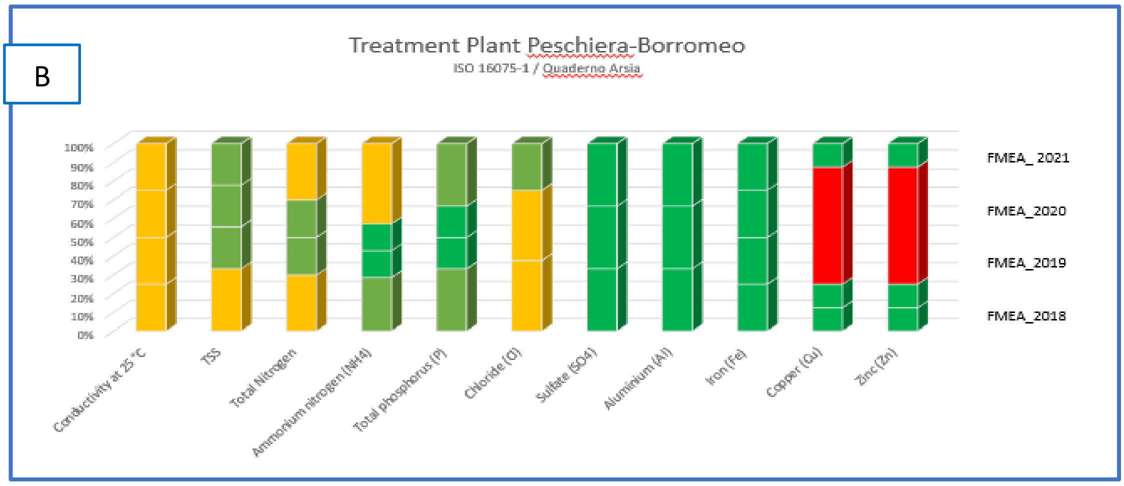


Figure 68 A and B: FMEA indices of investigated parameters representative of the quality of treated wastewater at Peschiera-Borromeo WWTP Line 2 for the years 2018, 2019, 2020 and 2021 considering the range [LOQ, ISO-L/ARSIA-L]

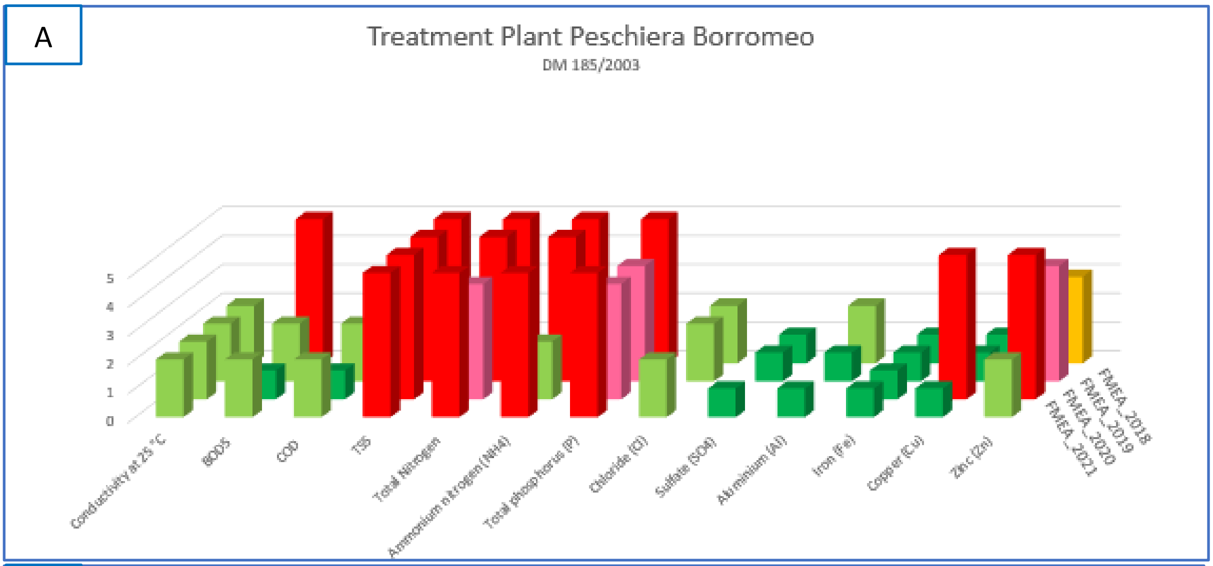
Finally, measured concentrations for the selected parameters in the wastewater effluent have never exceed the ISO-L limits. Indeed, the 95th percentile value is always less than 63 % of the

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

limits reported by ISO 16075-1:2020 for the most significant parameters, and consequently the accomplishment of a QCRA is not justified.

Results of the FMEA model using limits set by DM 185/2003

The histograms in Figure 69 A and B show the FMEA model applied to the same lab data set used for the elaboration shown in Figure 68, but obtained using different range of extremes [LOQ, LL].



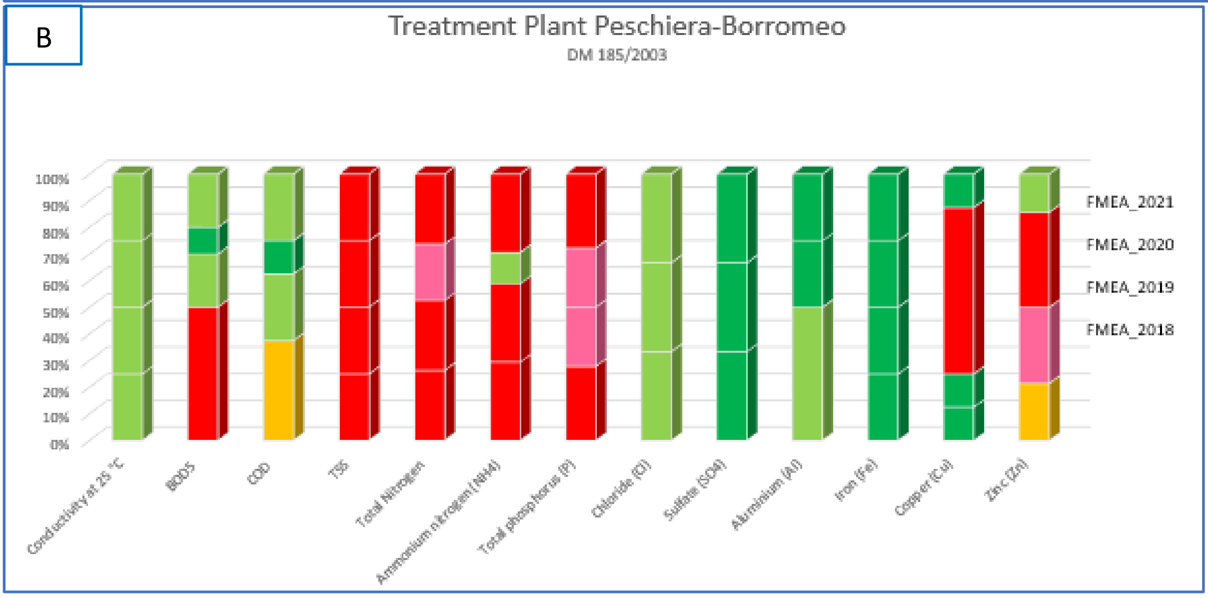


Figure 69 A and B: FMEA indices of investigated parameters representative of the quality of treated wastewater at Peschiera-Borromeo WWTP Line 2 for the years 2018, 2019, 2020 and 2021 considering the range [LOQ, DM 185/2003]

As it can be seen by the comparison of Figure 68 and Figure 69 , the limits imposed by the Italian DM 185/2003 are much more restrictive than those set by the ISO 16075-1:2020 and Quaderno Arsia limits.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

In fact, while in the previous analysis there were no parameters worthy of attention, in this case, according to Italian law limits, TSS, total nitrogen, ammonium nitrogen and total phosphorus parameters were found in class 5 in 2021. According to these results an intensification of the monitoring program could be justified for the previous cited parameters, since their values were found to be near the threshold (LL) of the national law 185/2003. However, the accomplishment of a QCRA is not justified, since none of the monitored parameters overcome the limits established by the National Regulation (LL) 185/2003.

Results of the Principal Components Analysis (PCA)

In order to confirm the results and to show the power of the employed approach, a *Principal Components Analysis* (PCA) was carried out for selected water quality parameters that were often detected at concentration levels above the limit of quantification.

Not of easy lecture, the Principal Component Analysis (PCA) is a technique for reducing the dimensionality of datasets, increasing interpretability but at the same time minimizing information loss. It does so by creating new uncorrelated variables that successively maximize the variance. Finding such new variables, the principal components analysis solves an eigenvalue/eigenvector problem.

In Figure 70, the results of the PCA analysis carried out for the selected parameters from 2018 (in blue) to 2021 (in purple) are reported. The plot was accomplished using the main three determined components, which explain largely the variance for all the variables.

The PCA biplot (Figure 70) is a type of plot where axes are the principal components and define the space of the variables. The projections of the data on the variable space, or samples scores, are represented by dots, whereas the blue lines represent variables loadings, or vectors that quantify the correlation of the variables with the components.

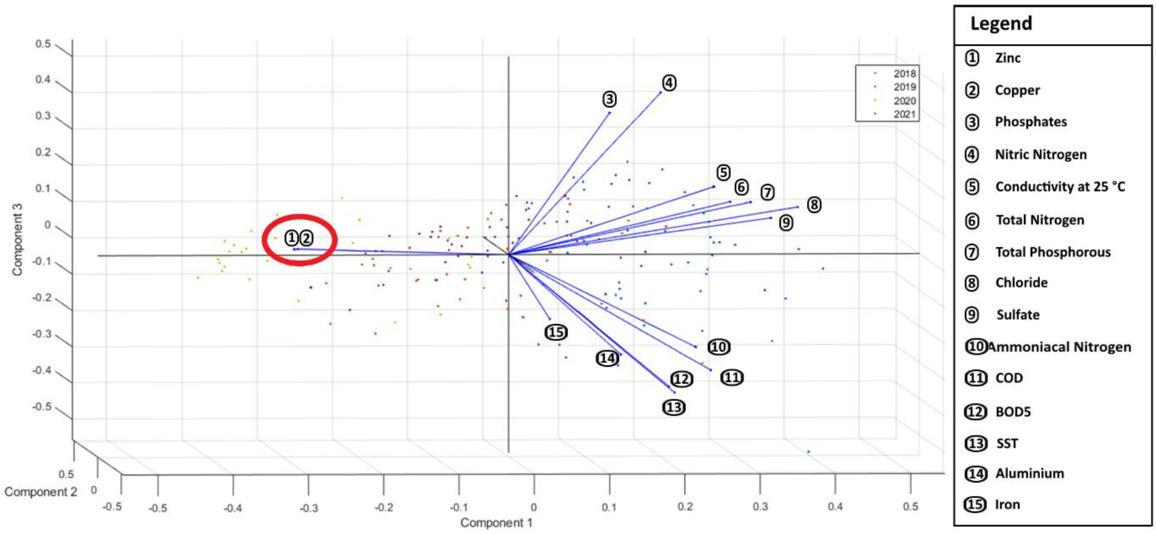


Figure 70: PCA biplot showing both PC scores of the samples (dots) and loading of variables (vectors)

In the PCA biplot, it is observed that parameters are well-separated in space. The first and second components explain the 60% of variance. The first component can explain the content of heavy metals in waters. In our study, metal profile in wastewater were evaluated through zinc, copper, aluminium, and iron concentrations, since the concentrations of the other heavy metals resulted often or always below their LOQ. In the same plot of Figure 70, it is possible to observe that, unlike from iron and aluminium, zinc and copper (red circle) represent a self-stand cluster. These two parameters are well correlated between them, but are not correlated with other parameters. This allows to conclude that the increase in concentration of these two metals observed in 2020 was a unique and random event.

The second component is strongly correlated with total nitrogen, total phosphorus, conductivity, BOD_5, COD and SST, and for this reason it can be seen as representative of the variation of the organic content of the system. Nitric nitrogen and phosphorus are positively correlated with the third component. The low variability of the data along the third component and the absence of correlation with BOD_5, COD and ammonium nitrogen demonstrates the normality of the trophic state of the waters, and consequently the good performance of the treatment plant.

Correlations found by PCA confirm the good performance of the WWTP in treating wastewater and the absence of high concentration of dangerous contaminants in the effluent during the observation periods. Hence the performance of a QCRA results not necessary.

3.3. Digital architecture of the Early Warning System at Peschiera-Borromeo WWTP

The purpose of Risk Management is to select, plan, establish and monitor risk reduction measures. In this context, the Early Warning System (EWS) can be used as a tool to support risk management, since it can receive real time information about effluent water quality from Peschiera Borromeo WWTP. In this way, decision making about water reuse can be supported and risk minimized allowing rapid reactions in case the occurrence of a hazardous event is detected by the EWS.

The EWS provides warnings if in the plant are detected conditions where quality requirements for water reuse cannot be satisfied. In these terms, the EWS can be used as a tool to support decision making. The rapid detection of a possible hazardous event needs an equally rapid communication to the strategy controllers of the water utility to let them promptly react.

Parameters of interest to be monitored by the EWS were selected considering the outcomes of the risk assessment (i.e., risk matrix, QMRA and QCRA) and the regulatory requirements for water reuse. The EWS architecture includes the generation of warning and alarms related to measurements obtained by on-line sensors and from machine learning algorithms (i.e., soft sensors). Particularly, for monitoring the microbial hazard the following device/sensors can be selected:

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

- ALERT System device for an almost real-time measurements of *E. coli* in the treated effluent;
- Sensors and soft sensors for the determination of TSS in the final effluent in real-time
- Soft sensors for the predictive determination of TSS in the final effluent (prediction of TSS concentration in the final effluent obtained up to 6 h earlier)
- Sensors for monitoring light intensity of UV lamps for disinfection

For monitoring the compliance of wastewater quality with the European Regulation 741/2020 on water reuse, the following sensors were selected:

- Soft sensor for the determination of BOD_5 in the final effluent in real-time
- Soft sensor for the determination of COD in the final effluent in real-time
- Sensors for real-time measurements of TOC and UV absorbance at 254 nm, which can be used to estimate COD and BOD_5.

Hence the warning/instruction messages listed in Table 30 can be generated.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Table 30: Warning messages of the EWS for water reuse at Peschiera Borromeo WWTP

Sensor	Parameter of interest	Threshold value	Warning/instruction message	Note
Alert System	E. Coli	10 (number/100 mL)	Water quality non-suitable for class A reuse	No compliance with class A water quality
Alert System	E. Coli	100 (number/100 mL)	Water quality non-suitable for class B reuse	No compliance with class B water quality
Alert System	E. Coli	1000 (number/100 mL)	Water quality non-suitable for class C reuse	No compliance with class C water quality
Soft sensor (including forecasting measurements)	TSS	10 mg/L	Stop water reuse	High TSS concentration in the final effluent do not allow a proper disinfection by the UV lamps
Sensor	UV light intensity	Sudden drop of UV light intensity	Stop water reuse	Proper UV disinfection is not ensured
Soft sensor/ UV absorbance sensor	BOD_5	10 mg/L	Water quality non-suitable for class A reuse	No compliance with class A water quality
Soft sensor/UV absorbance sensor	BOD_5	25 mg/L	Stop water reuse	No compliance with EU regulation 741/2022 on water reuse
Soft sensor/UV absorbance sensor/TOC	COD	125 mg/L	Stop water reuse	No compliance with EU regulation 741/2022 on water reuse

Since microbial risk was highlighted as one of the most significant risk during wastewater reuse, in this study, particular attention has been addressed to TSS prediction since solid particles may be vehicle of pathogens. Particularly, from data collected in DWC project during

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

sampling campaigns to test Alert Systems, it was observed that lab TSS and *E. coli* concentration can be correlated, as shown in Figure 71.

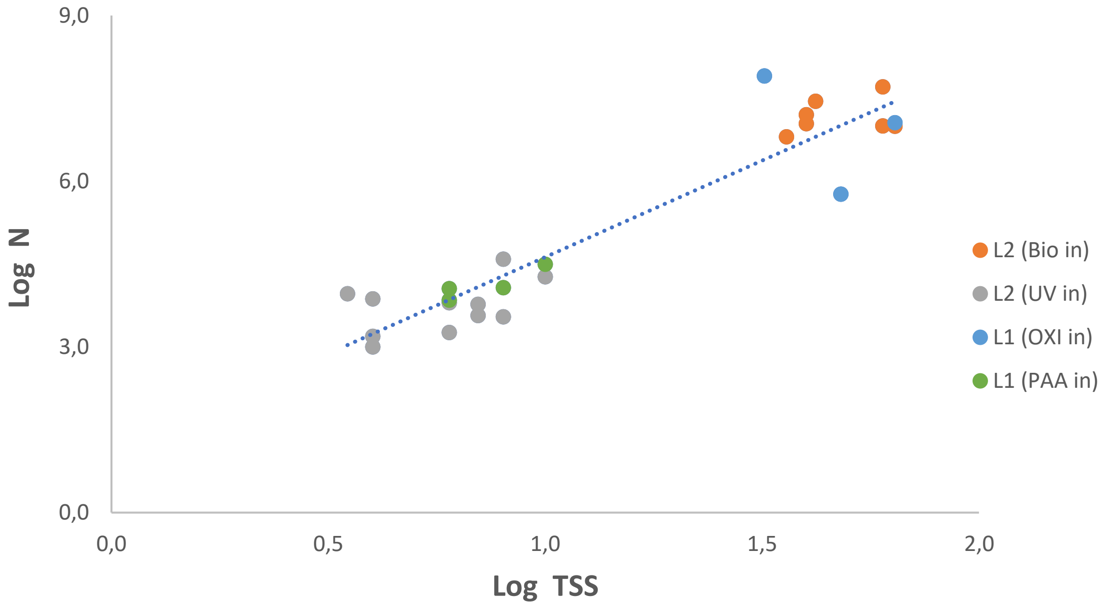


Figure 71: Correlation between E.coli (expressed as LogN) and TSS (expressed as LogTSS) lab data at different treatment stage: influent to biologic unit of Line 2 (orange); Influent to UV unit of Line 2 (grey); influent to biologic unit of Line 1(blue); influent to disinfection unit of Line 1 (green)

Moreover, high TSS concentration can cause a screen effect and a subsequent decrease on disinfection performances of the UV disinfection lamps. At Peschiera Borromeo WWTP, TSS concentration in the final effluent is generally lower than 10 mg/L. Hence, this value was considered as the threshold value for TSS concentration in the wastewater effluent that can assure a proper disinfection by UV lamps. In the case of TSS concentration higher than 10 mg/L, water reuse should be stopped since a satisfactory disinfection is not assured.

Resuming, the Early Warning System has been developed at Peschiera Borromeo WWTP to forecast water quality depending on sensors measurements and soft-sensing techniques. Artificial Neural Networks (ANN) were developed for real-time and time-series prediction of parameters related to wastewater quality. According to the WWTP operational conditions, the soft-sensing algorithms are able to forecast the quantitative values of target parameters and compare them with threshold limits for water reuse. The developed ANNs are able to predict the following parameters:

- Biochemical oxygen demand (BOD_5 – real-time prediction);
- Chemical oxygen demand (COD – real-time prediction);
- Total suspended solids (TSS – real-time prediction and up to 6 hours earlier prediction).

Soft sensors are smart process models that use “easy-to-measure” variables, which can be monitored on-line at a reasonable cost, to predict target “hard-to-measure” variables (F. A. A.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Souza et al., 2016). The easy-to-measure or secondary parameters are typically pressure, temperature, flow rate, pH, conductivity, turbidity, and dissolved oxygen (i.e., typical operational parameters in a WWTP).

Real-time monitoring of BOD_5 and COD in WWTP can be accomplished indirectly by surrogate measurements, such UV absorbance at 254 nm or by estimation from TOC on-line measurements. Specific sensors are available in the market to predict TSS.

However, TSS, BOD, COD were predicted by soft-sensing technique since the sensors installed at Peschiera Borromeo WWTP were not able to produce reliable measurements. In an ideal case, developed soft-sensors may support real sensor during maintenance/malfunction periods.

Periodic lab analyses are anyway scheduled for standard monitoring programs and are used to guarantee the precision of soft-sensors estimations.

Furthermore, for the parameter TSS a time series ANN was developed to predict earlier in time (up to 6 hours earlier) the TSS concentration in the final effluent and give the possibility to the operational facility staff to stop water reuse or take preventive actions before the occurrence of the hazardous event.

As it was largely discussed in Deliverable 1.1, sensors for the measurements of water quality parameters did not produce reliable data at Peschiera Borromeo WWTP (particularly the TSS probe). In addition, very few laboratory measurements were available for BOD_5, COD and TSS. Since a very large set of data is needed for the development of high-performing ANN, a digital twin model of the WWTP was created using the software BIOWIN. Hence, data generated by software simulation were used to train and test the predictive artificial neural networks. Finally, the developed ANNs were validated with real data by a domain adaptation procedure. Indeed, this procedure allowed to develop a soft sensor able to provide predicted concentrations of TSS, COD and BOD5for real case application using a limited set of real data.

Even if data from the new installed sensors at Peschiera Borromeo WWTP were not available in time for the submission of the Deliverable D1.1, recently signals from TOC sensor have been connected and stored in remote control. Raw data (Figure 72) from TOC sensor are provided every 40 minutes, and they report also an estimation of COD concentrations, which is calculated by using an internal conversion factor. However, since TOC data were limited in number and their reliability need to be checked by comparison with laboratory data, they were not used for soft-sensors development.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

TIME	DATE		CO2z	TICmgu	TICmgc	CO2p	TOCmgu	TOCmgc	CO2p	CODmgO/l	BT_DegC	MB_DegC	Atm
12:19:43	24-03-22	S1:1	-7.56	42.75	46.22	1502.0	6.77	2.77	227.5	8.87	33.2	39.0	101.8
11:28:16	24-03-22	S1:1	-7.47	43.03	46.16	1549.5	6.35	2.91	226.8	9.32	32.9	38.0	101.8
10:36:50	24-03-22	S1:1	-7.29	42.37	45.93	1486.6	6.67	3.00	242.0	9.59	32.0	38.0	101.9
09:45:23	24-03-22	S1:1	-7.37	42.66	46.01	1464.8	7.24	3.08	237.6	9.85	31.2	36.0	101.8
08:52:56	24-03-22	S1:1	-7.43	42.51	45.90	1421.6	6.64	2.95	222.3	9.44	29.5	35.0	101.8
08:01:30	24-03-22	S1:1	-7.71	46.77	51.16	1545.7	6.95	2.98	241.8	9.55	26.3	31.0	101.8
07:10:03	24-03-22	S1:1	-7.68	46.26	52.29	1597.9	6.83	2.81	243.0	9.01	24.8	30.0	101.8
06:18:37	24-03-22	S1:1	-7.67	49.55	53.26	1648.2	6.76	2.80	221.4	8.97	24.9	31.0	101.8
05:27:10	24-03-22	S1:1	-7.55	49.95	53.67	1692.8	6.38	2.67	219.5	8.53	25.6	31.0	101.8
04:35:44	24-03-22	S1:1	-7.50	48.92	53.62	1673.6	6.80	2.71	229.5	8.69	26.0	32.0	101.8
03:44:17	24-03-22	S1:1	-7.53	50.64	54.56	1729.2	6.29	2.49	213.4	7.97	26.5	32.0	101.8
02:52:50	24-03-22	S1:1	-7.68	49.62	53.49	1712.8	7.46	3.54	233.0	11.33	27.2	33.0	101.8
02:01:24	24-03-22	S1:1	-7.89	49.65	53.54	1665.2	8.74	4.69	265.5	15.00	27.9	34.0	101.8
01:09:57	24-03-22	S1:1	-8.17	48.96	51.80	1645.1	7.64	3.33	266.9	10.65	28.5	34.0	101.9
00:18:31	24-03-22	S1:1	-8.29	48.99	50.63	1678.1	7.37	3.38	243.7	10.80	29.3	35.0	101.9

Figure 72: Extract of raw data acquired from TOC sensor

3.3.1. Modelling of wastewater treatment processes

Predictive models of the Early Warning System require an affordable and large set of data to be trained and validated, including both periods of normal operation and periods with malfunctions occurrence, which are rarely available. Hence, a model representing typical operational conditions of Peschiera Borromeo WWTP was created using the simulation software BIOWIN. Particularly, models were used to obtain simulated data representative of the measurements performed by sensors that are actually installed at the investigated WWTP.

Each process unit of Peschiera Borromeo WWTP was schematized using dimensional data as well as information about typical working conditions.

In detail, the SEDIPAC unit was schematized through a Lamellar sedimentation unit using data of its volumetry, removal efficiency and sludge underflow. The BIOFOR unit was modelled using a Bioreactor – Media unit. In detail, for the schematization of the anoxic compartment were provided information about size of the tanks and percentage of volume filled with media to support biomass growth. For the aerobic tanks additional data about diffuser systems and air flowrates were provided (Figure 72).

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

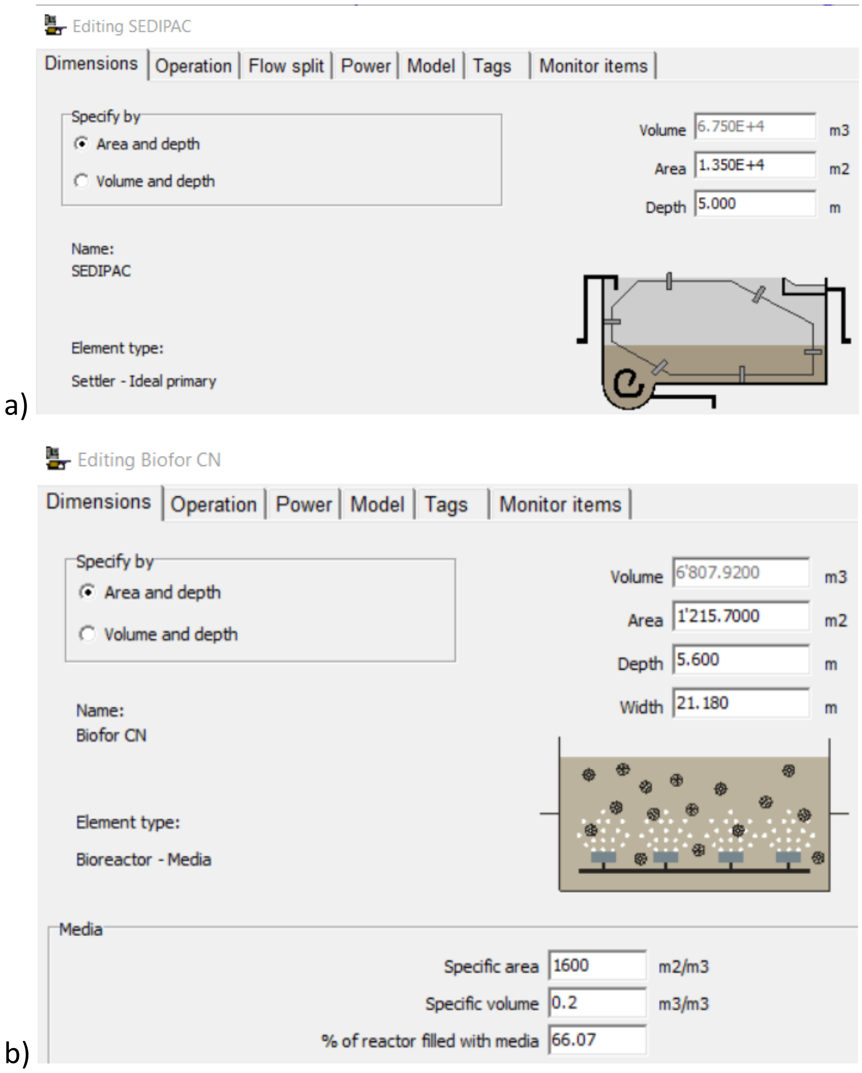


Figure 72: BIOWIN schematization of a) SEDIPAC and b) BIOFOR aerobic units

Data on internal mixer liquor recirculation and backwash were included to simulate WWTP operational conditions. Data of the dosage of external carbon source for denitrification and of chemicals for phosphorus precipitation were also included. The layout created by BIOWIN for the simulation of the wastewater treatment processes at Peschiera Borromeo WWTP is shown in Figure 73.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

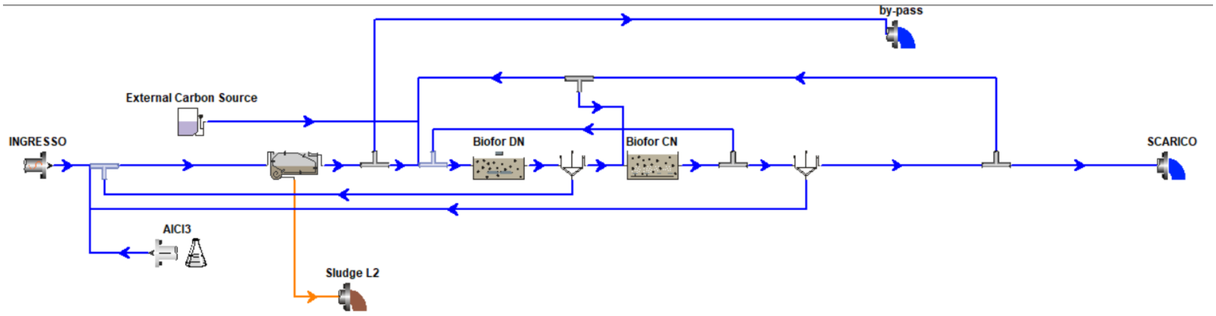


Figure 73: Peschiera Borromeo WWTP layout

Once the configuration was completed, the model was run using real data coming from the WWTP. Laboratory data and sensors data measured during the whole 2020 were used for the purpose. Data were provided as hourly parameters and included:

- Influent flowrate;

- influent COD

- Influent pH;

- Influent TKN (Total Kyendal Nitrogen);

- Influent TP (total phosphorus);

- Temperature in the BIOFOR reactor;

Data for temperature and pH were available from sensors with a 15 minutes time steps. Hourly averages were calculated to have hourly values. Data for flowrate, COD, TKN, TP were available as daily averages because related to laboratory analyses of 24 h composite samples (except flowrate data, which were obtained daily by a totalizator device). Hence, the generation of hourly variations for the above-mentioned parameters was accomplished using the tool developed by Langergraber et al., (2008), which allows to simulate typical diurnal variation of wastewater quality parameters. Laboratory data were available to cover different seasonal conditions within one year. Hence, hourly data were generated for one year of operation of Peschiera Borromeo WWTP. Static simulations by BIOWIN were performed to calibrate and validate the model by comparing simulated data with available laboratory data. On the contrary, dynamic simulations of the software BIOWIN were performed to generate a data set with 1 hour time steps that cover one year period for the following parameters:

- Dissolved oxygen concentration in the BIOFOR reactor

- Effluent flow rate;

- Effluent pH;

- Effluent total Nitrogen (TN)

- Effluent N-NH_4;

- Effluent N-NO_x;

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

- Effluent $P\text{-}PO_4$

- Effluent COD

- Effluent TSS

- Effluent BOD_5

In Figure 74 are reported, as an example, obtained concentrations for COD and TSS in the wastewater effluent during the simulated year 2020.

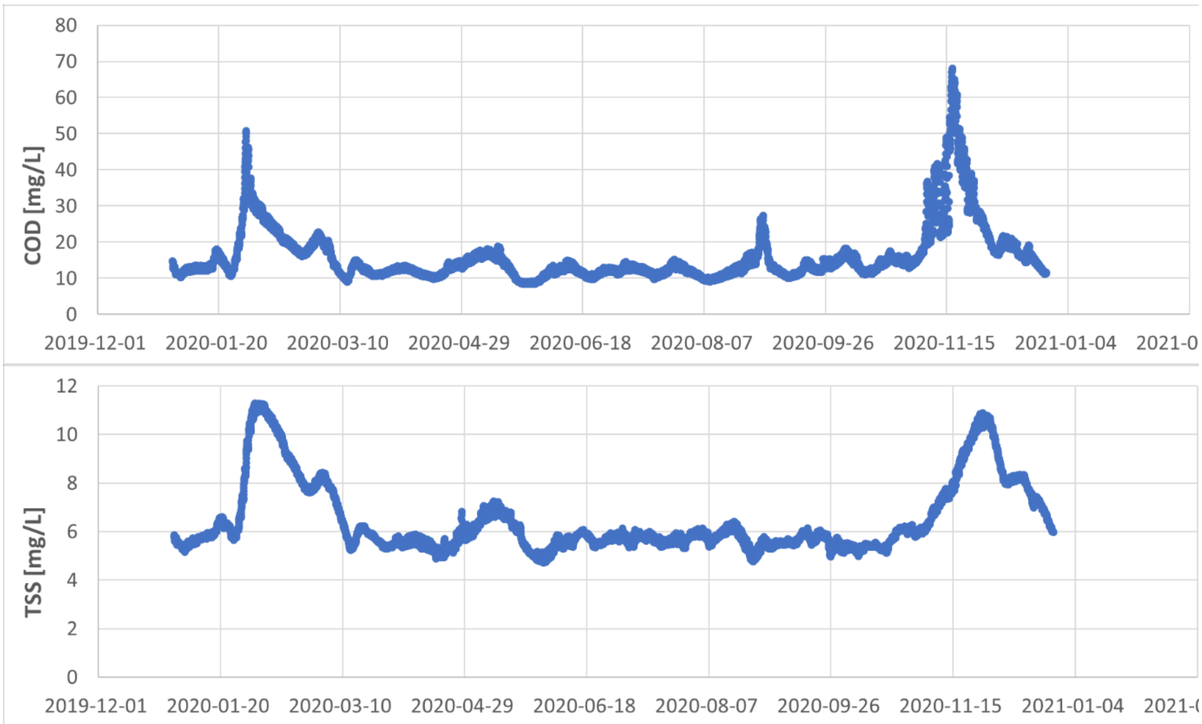
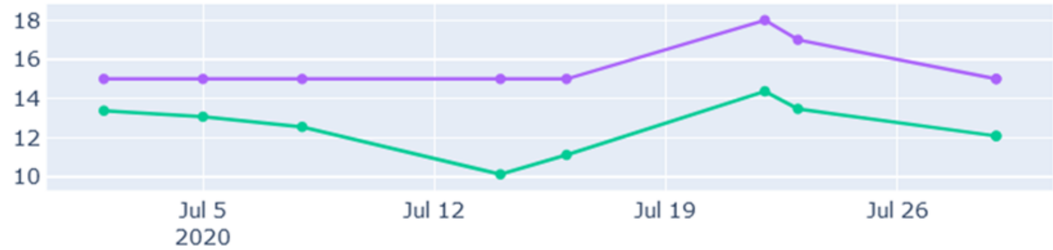


Figure 74: Obtained hourly data concentrations for COD and TSS by BIOWIN simulation

In Figure 75 is reported a comparison between laboratory data and calculated daily averages of data obtained by BIOWIN simulation for COD and TN in July 2020. The comparison shows a good agreement between simulated and real data. It must be highlighted that the graph in Figure 75 is referred to a limited period, while a longer dataset has been considered for model validation. In any case, the obtained average error between simulated and real data was kept within the 10 – 20 %. In addition, the analytical methods for COD determination are characterised by an uncertainty of around 5 – 10 mg/L, which must be taken into account when comparing lab data with simulation results.

COD

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.



in purple lab data, in green BIOWIN data

TN

in green lab data, in red BIOWIN data

Figure 75: Comparison between real and simulated data for the parameters COD and TN

To increase the dataset and to simulate conditions, where the investigated WWTP is not working properly, different malfunction scenarios were simulated by BIOWIN. Simulated malfunction events with indication of the modified operational parameters are reassumed in Table 31. Malfunctions were simulated in both warm and cold seasons to take into account scenarios with different temperature of operation.

Table 31: Malfunction scenarios simulated by BIOWIN

Malfunction of the Aeration system

Simulated event	Aeration interruption (breakage of the blowers)	Insufficient aeration (reduction of air flow rate)
Modified parameter during simulation	Q air (air flowrate)	Q air
Main affected parameters	DO in BIOFOR reactor, NH4, NOx, COD, BOD, TSS in the effluent	DO in BIOFOR reactor, NH4, NOx, COD, BOD, TSS in the effluent
Simulated Scenarios	Q air = 0 m3/h for different time intervals	50% reduction of Qair for different time intervals
scenario_1	Qair = 0 for 1 h	Qair = -50% for 1 h
scenario_2	Qair = 0 for 5 h	Qair = -50% for 5 h
scenario_3	Qair = 0 for 12 h	Qair = -50% for 12 h
scenario_4	Qair = 0 for 24 h	Qair = -50% for 24 h
scenario_5	Qair = 0 for 48 h	Qair = -50% for 48 h

Error in sludge underflow management

Simulated event	Increase of sludge underflow	Reduction of sludge underflow

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Modified parameter during simulation	Q underflow in the SEDIPAC unit	Q underflow in the SEDIPAC unit
Main affected parameters	TSS in the effluent	TSS in the effluent
Scenarios	increase of Q underflow	reduction of Q underflow
scenario_1	Q underflow +20% for 2 days	Q underflow -20% for 2 days
scenario_2	Q underflow +50% for 2 days	Q underflow -50% for 2 days

Internal recycles

Simulated event	Errors in setting the Internal recycle of sludge to the SEDIPAC unit	Malfunction of the backwash system
Modified parameter during simulation	Q of sludge recycle	Q backwash
Main affected parameters	TSS, NH4, NOx, COD in the effluent	TSS, NH4, NOx, COD in the effluent
Scenarios	Internal Recycle variations	variation of Q backwash
scenario_1	Q r -20% for 1 day	Q backwash +30% for 2 days
scenario_2	Q r +20% for 1 day	Q backwash -30% for 2 days
scenario_3	Q r = 0 for 1 day	Q backwash = 0 for 2 days

Industrial discharge

Simulated event	Industrial discharge	Industrial discharge
Modified parameter during simulation	pH	pH
Main affected parameters	NNH4, NNOx, COD, BOD in the effluent	NNH4, NNOx, COD, BOD in the effluent
Scenarios	Influent with pH = 5	Influent with pH = 11
scenario_1	pH = 5 for 1 h	pH = 11 for 1 h
scenario_2	pH = 5 for 5 h	pH = 11 for 5 h
scenario_3	pH = 5 for 12 h	pH = 11 for 12 h
scenario_4	pH = 5 for 24 h	pH = 11 for 24 h
scenario_5	pH = 5 for 48 h	pH = 11 for 48 h

High organic or hydraulic load

Simulated event	Organic overload (es., discharge of agri-food industry)	Rain event
Modified parameter during simulation	COD in the influent	Q, COD, TKN in the influent
Main affected parameters	DO in the BIOFOR, COD, NH4, NOx, BOD, TSS in the effluent	DO in the BIOFOR, COD, NH4, NOx, BOD, TSS in the effluent

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Scenarios	Increase of COD load	Increase of influent Q flowrate
scenario_1	influent organic load equal to the 50% of maximum COD measured in the influent of the plant for 1 day	Q max, COD min, TKN min measured in the influent to the plant for 1 day
scenario_2	influent organic load equal to the maximum COD measured in the influent of the plant for 1 day	-

Chemical dosing

Simulated event	Errors in external carbon dosing	Errors in AlCl3 dosing for phosphorus removal
Modified parameter during simulation	Q of dosed external carbon	Q of dosed chemicals
Main affected parameters	COD, NNOx in the effluent	PPO4, TP in the effluent
Scenarios	variation of Q dosed	variation of Q dosed
scenario_1	Q -50% for 1 day	Q -20% for 1 day
scenario_2	Q +50% for 1 day	Q +20% for 1 day

Figure 76 shows the BOD_5 concentration in the effluent during the BIOWIN simulation of malfunction of the aeration system.

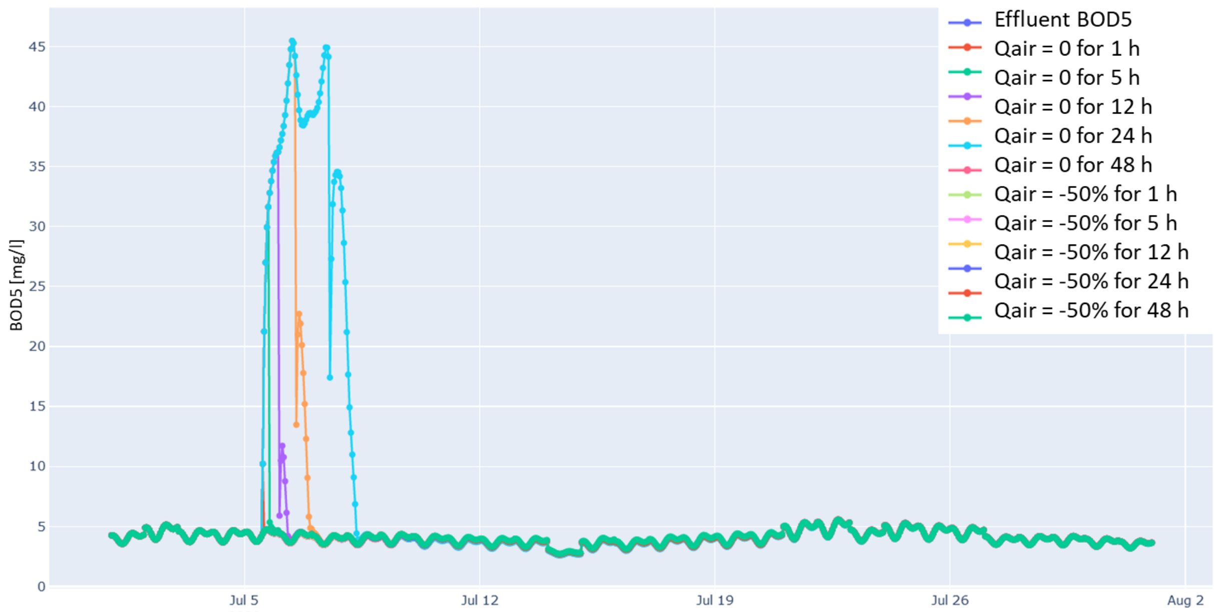


Figure 76: BOD_5 concentration in the effluent during BIOWIN simulation of malfunction related to the incorrect operation of the aeration system in the BIOFOR reactor.

Obtained hourly data were used to develop the soft sensors included in the EWS to generate the warnings/alarms reassumed in Table 30. The complete dataset included 10330 data points for each selected parameter.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

3.3.2. Soft-sensors development for wastewater quality prediction

As explained, the obtained database by BIOWIN simulation was used to feed and train two different Artificial Neural Networks (ANN):

1. The first soft sensor was developed by creating an ANN able to predict TSS, BOD, COD from parameters that are continuously monitored by sensors at the plant
2. The second soft sensor was developed to predict TSS concentration up to 6 hours earlier in time to give technicians the possibility to stop water reuse or take preventive actions before the occurrence of the hazardous events that may affect the safety of the water reuse.

Since microbial hazards and disinfection performances are related with solid concentration, both ANNs were developed to obtain, as output, the TSS concentration at the effluent. Furthermore, BOD and COD were predicted by the first ANN since they are regulated parameters for wastewater reuse.

3.3.2.1. Development of soft sensor for predicting TSS, COD, BOD

The developed network used for real-time prediction of TSS, COD, BOD was a Deep Feed Forward Neural Network (DFF) that taking as inputs (simulated) sensors data could provide predictive trends of not-measured parameters (i.e., TSS, COD or BOD). The ANN was performed using 6 inputs, including Influent flowrate, Influent pH, temperature in the anoxic BIOFOR reactor, dissolved oxygen (DO) concentration in aerated BIOFOR reactor, effluent flowrate, effluent pH. These parameters were selected as input of the ANN because they are easy-to-measure by cheap and reliable sensors in a WWTP, and they are the parameter generally used for process control during wastewater treatment. Furthermore, many WWTPs are equipped by these sensors, and, thus, the developed ANN may be adapted and replicated in other WWTPs very easily.

The developed ANN in this study was a MULTITASK (MLT) model able to predict concurrently three target parameters (i.e., TSS, BOD, COD). Indeed, in the context where more tasks for prediction are needed, which rely on a common hidden layer representation, a MLT approach can provide the following benefits (Ruder, 2017):

- MTL represents an implicit data augmentation strategy which allows to avoid model overfitting. As different tasks could have different noise patterns, a model that learns more tasks simultaneously is able to learn a more general representation. Learning just one task independently bears the risk of overfitting to that specific task, while learning all the tasks jointly enables the model to obtain a better representation function through averaging the noise patterns.
- If a task is very noisy or data is limited and high-dimensional, it can be difficult for a model to differentiate between relevant and irrelevant features. MTL approach can help the model to focus its attention on those features that matter as other tasks will provide additional evidence for the relevance or irrelevance of those features.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

- MLT speeds up the inference phase: on a real use-case setup, it is faster to obtain all the predictions from one model respect to three. Moreover, the computational cost of training just one model is lower than training three independent models.

The architecture of the selected MTL model is constituted by the following layers (Figure 77):

- Input layer
- Shared layer (Dense layer with 32 neurons)
- Task specific layer (2 Dense layers per task, with 16 and 8 neurons respectively)
- Output layer (1 single neuron per task)

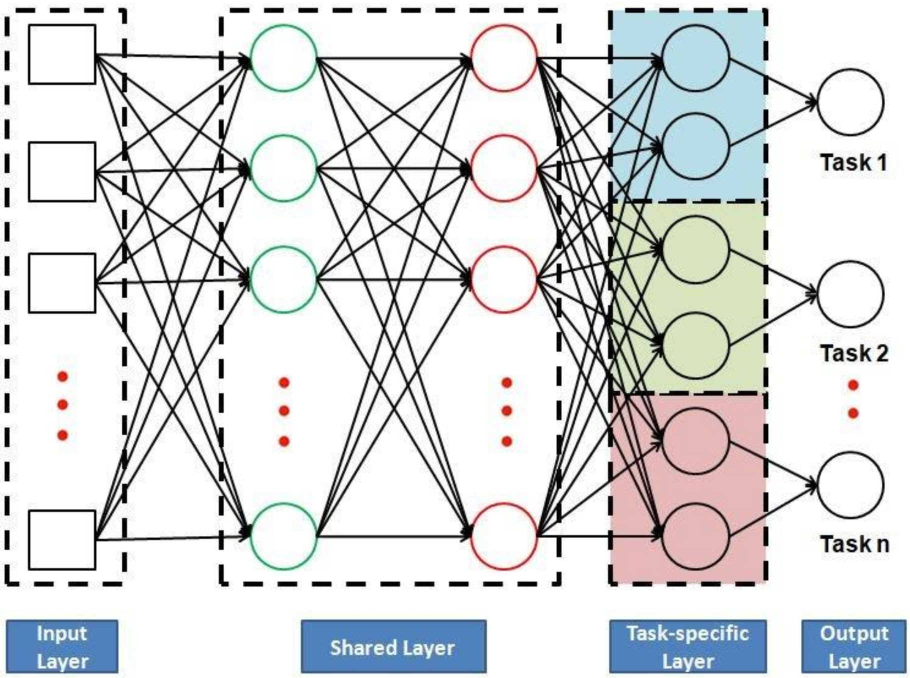


Figure 77: Schematic representation of the selected multitask ANN model.

The ReLU Function was selected as the transfer (activation) function utilized by each neuron to generate its output. Hence the ANN was developed selecting 70% of BIOWIN simulated data for training, 10% for validation, and 20% for testing (Figure 78). The training and testing samples were randomly selected from the full dataset.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

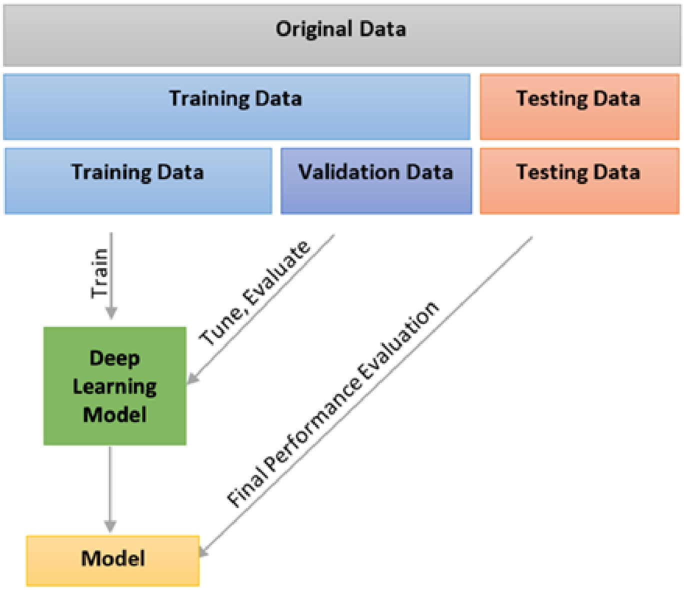


Figure 78: Schematic representation of the development process of an ANN model

To evaluate the performance of the developed ANN models, statistical parameters including Correlation Coefficient (CC), Root Mean Square Error ($RMSE$), Scatter Index (SI), $BIAS$, Mean of Absolute Percentage Difference *(MeanAPD)* and its standard deviation *(StdAPD)* were calculated by the following equations.

$$CC = \frac{\sum_{i=1}^{N}(S_i - \overline{S_m})(P_i - \overline{P_m})}{\sqrt{\sum_{i=1}^{N}(S_i - \overline{S_m})^2 \times \sum_{i=1}^{N}(P_i - \overline{P_m})^2}} \quad (3.10)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(P_i - S_i)^2}{N}} \quad (3.11)$$

$$SI = \frac{RMSE}{\overline{S_m}} \times 100 \quad (3.12)$$

$$BIAS = \frac{\sum_{i=1}^{N}(P_i - S_i)}{N} \quad (3.13)$$

$$\mathrm{APDi} = \left|\frac{\mathrm{P_i} - \mathrm{S_i}}{\mathrm{S_i}}\right| \times 100 \quad \text{in } (\%) \quad (3.14)$$

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

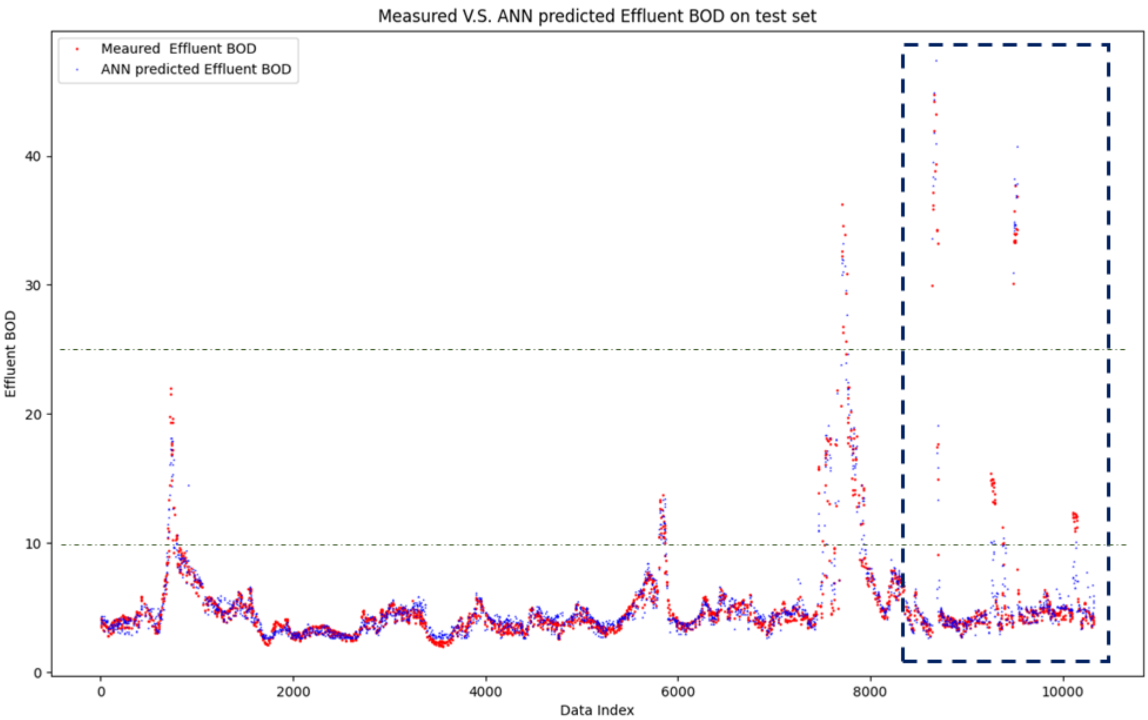
Where S_i and P_i denote the hourly-averaged Biowin simulations and their correspondent ANN predictions at i^{th} hour, respectively. N is the total number of available hourly values. $\overline{S_m}$ and $\overline{P_m}$ are mean values of hourly BIOWIN simulations and ANN predicted values, respectively. *MaenADP* and *StdADP* were obtained by calculating the average and standard deviation of all determined *APDi* (Absolute Percentage Differences*)* by equation 3.14.

For the testing phase, which represents the final evaluation of the performance of the developed ANN model, the statistic parameters above described are reported in Table 32.

Table 32: Statistical parameters calculated to evaluate the performance of the developed ANN model during the testing phase

Predicted parameter	CC	RMSE	SI	BIAS	MeanAPD (%)	StdAPD
Effluent BOD	0.963	1.235	0.228	-0.006	8.460	10.480
Effluent COD	0.961	2.388	0.147	-0.050	6.671	8.521
Effluent TSS	0.856	1.419	0.217	-0.105	5.630	11.966

In Figure 79 - Figure 82 is reported a comparison between the predicted data by the ANN models and the simulated BIOWIN data, which represent the true values, during the test phase. Overall, the calculated parameters in Table 32 confirm that the developed soft sensor was able to predict COD, BOD and TSS concentrations in the final effluent with a very good accuracy given both correlation (CC) and error indices (RMSE, SI, BIAS and MeanAPD).



digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Figure 79: Comparison between BOD hourly data predicted by ANN and simulated by BIOWIN. Data included within the rectangle in dashed line are related to the simulation of malfunction scenarios. Horizontal dashed lines identify thresholds for warning messages generation to stop water reuse

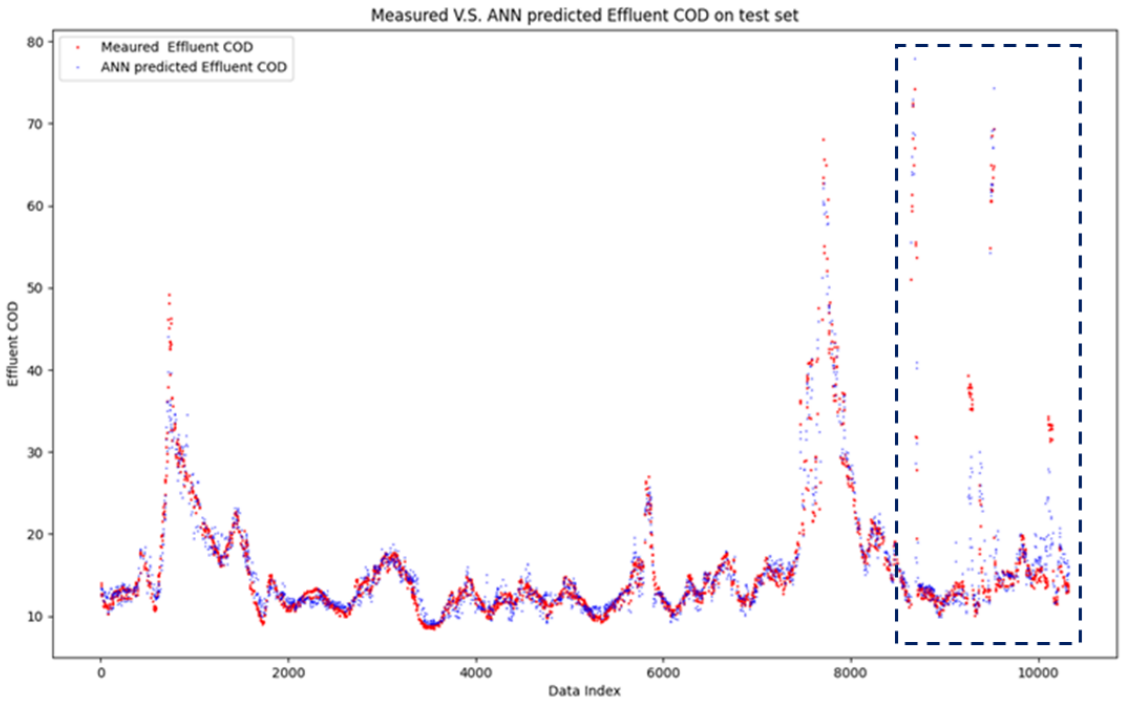
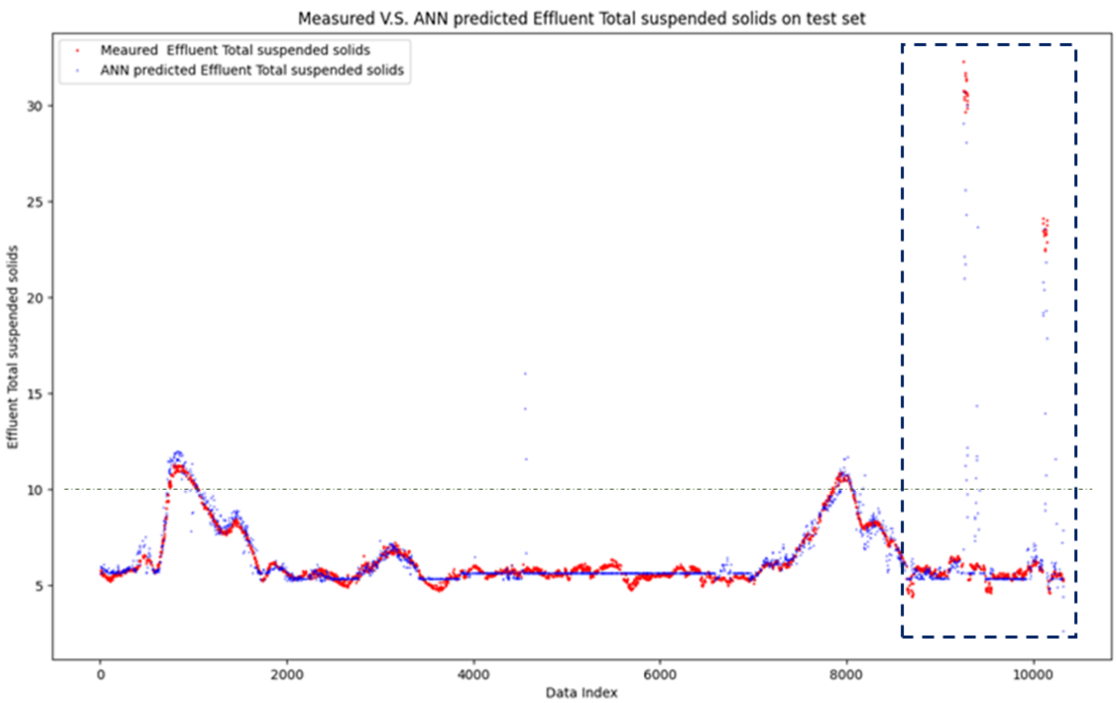


Figure 80: Comparison between COD hourly data predicted by ANN and simulated by BIOWIN. Data included within the rectangle in dashed line are related to the simulation of malfunction scenarios



digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Figure 81: Comparison between TSS hourly data predicted by ANN and simulated by BIOWIN. Data included within the rectangle in dashed line are related to the simulation of malfunction scenario. Horizontal dashed lines identify thresholds for warning messages generation to stop water reuse

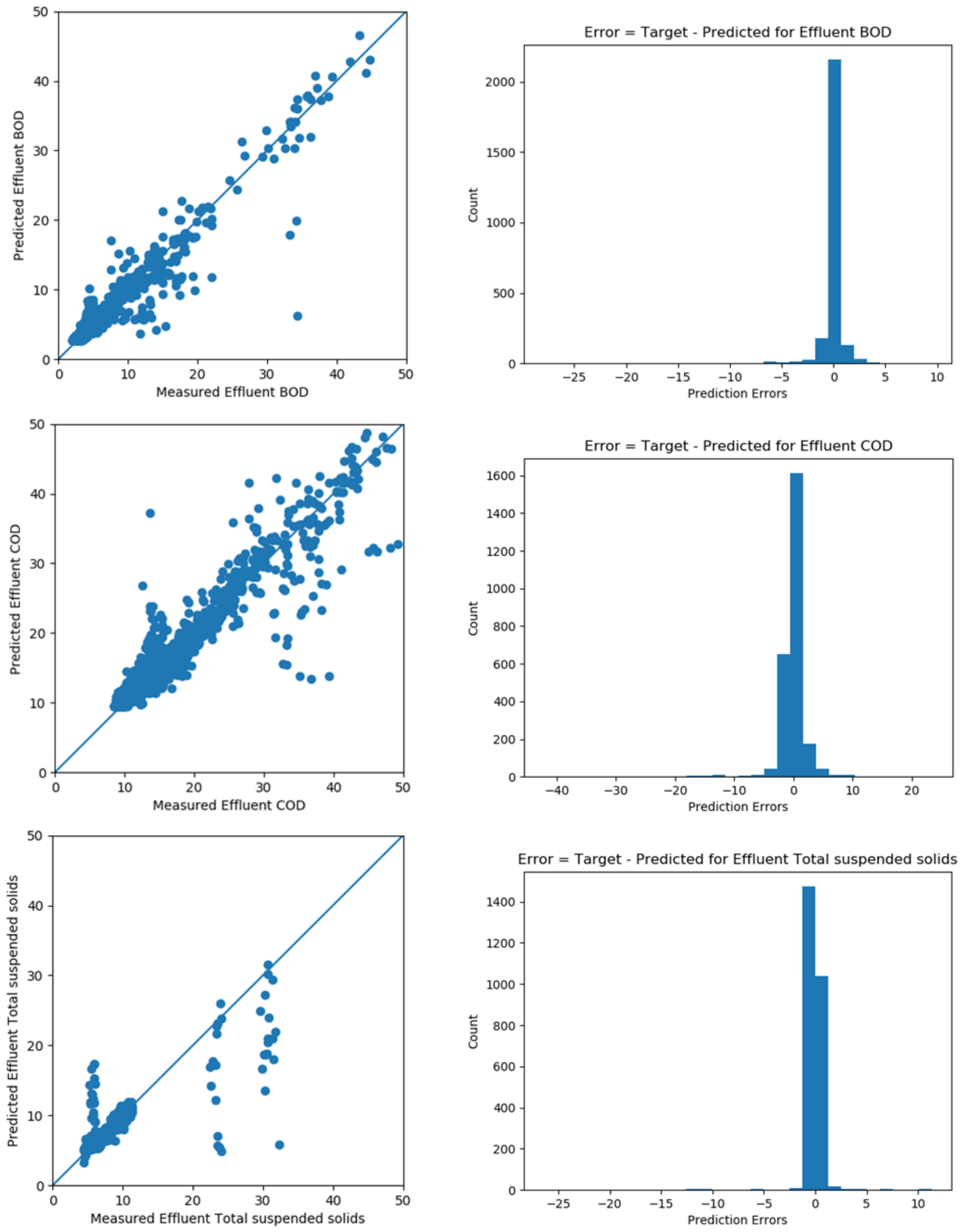


Figure 82: Scatter plots and errors histogram to compare hourly data predicted by ANN and simulated by BIOWIN

Nevertheless, an accurate analysis of Figure 79 - Figure 82 shows that the biggest deviation of predicted data from the true data occurred for events related to the simulated malfunction

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

scenarios (highlighted by a rectangle with dashed line in Figure 79 - Figure 81). It is not a surprising outcome, since malfunctions were simulated in BIOWIN by improvise changes of operating conditions, which created a sudden jump of the values for BOD, COD and TSS in the effluent. However, even though for these data points the error was a bit higher than for the rest of the dataset, it can still be affirmed that the developed soft sensor was able to detect the occurrence of the simulated hazard event, and the sudden change of BOD, COD and TSS values. Hence, it was able to absolve his function within the EWS since the warning message to stop the water reuse would have been generated by the algorithm. On the contrary, during normal operating conditions of the WWTP the soft sensor is always able to predict the target COD, BOD, TSS parameters with very high accuracy. Finally, it has to be noted that thresholds for water reuse shown in Table 30 were also exceeded during the simulated "normal" year of operation at Peschiera Borromeo WWTP by BIOWIN, and not only during malfunction scenarios. In these cases, the overcoming of the threshold values occurred gradually following a defined trend, and predicted data by ANN showed again a very high accuracy (Figure 79 and Figure 81).

3.3.2.2. Domain adaptation of the ANN model with real data

Since the data used for developing the ANN were generated via BIOWIN simulations and the real prob data were not exploited to train the ANN model, it was elaborated a procedure to adapt the configuration of the developed ANN model using a real data set. In this project, accurate prob measurements for all input and output parameters needed to train the ANN were not available. Hence, it was needed to combine prob records with lab measurements to prepare a real validation dataset. However, the produced dataset was still limited in number of datapoints. Totally, 1855 data points were extracted from both lab and prob records. Particularly, data of pH influent, pH effluent, DO in aerobic BIOFOR reactor and temperature in anoxic BIOFOR reactor were available by sensors and it was easy to obtain hourly data for each of these parameters. For influent and effluent flowrates (obtained by totalizator devices), and for BOD, COD, TSS concentrations in the effluent (available by laboratory measurements), data were available only as daily average values. Hourly variation during a day of the influent flowrate was obtained using the tool developed by Langergraber et al., (2008) described previously. For effluent flowrate, BOD, COD and TSS concentration in the effluent, it was assumed that hourly data had the same value of the correspondent daily averaged value. This assumption can be in part justified considering the equalization effect on wastewater quality parameters of the effluent obtained due to the high hydraulic retention time (HRT) of the WWTP. Further limitation of the available dataset was that laboratory data were limited and related to non-consecutive days.

In this project a domain adaptation procedure (Farahani et al., 2021) of the developed ANN was accomplished by feeding the model with the new field data set, and by using a cross-validation method (Bergmeir & Benítez, 2012). Particularly, 10 different iterations were performed splitting the total dataset in days (i.e., hourly data present within one day were not divided to make the ANN development procedure more suitable for real case applications),

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

and using 50% of available days for training, 10% for validation, and 40% for testing (Figure 83). In this way, for each iteration were available:

- Training: 40 days (1000 data points)
- Validation: 9 days (115 data points)
- Test: 34 days (740 data points)

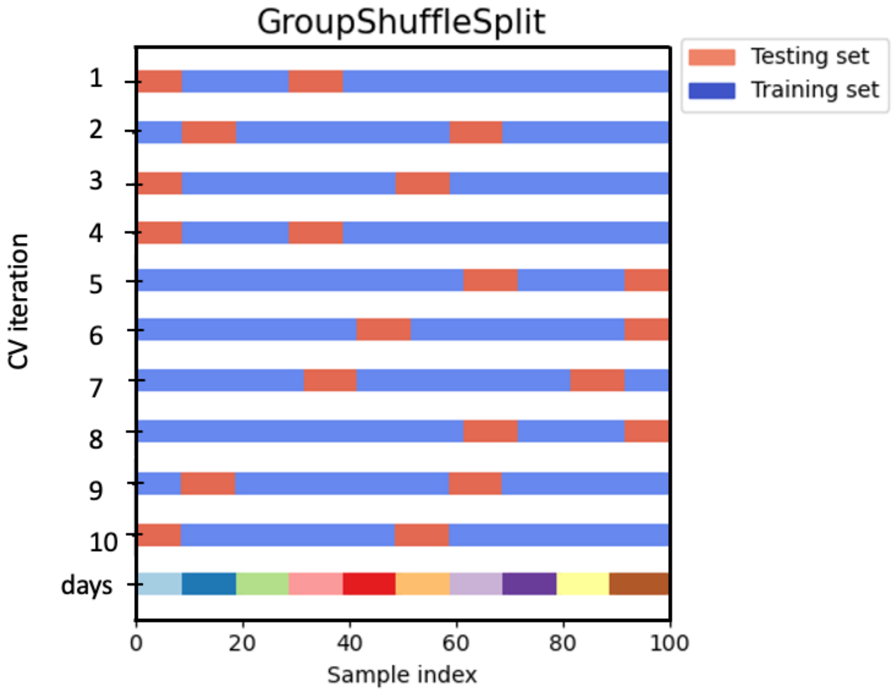


Figure 83: Group Shuffle split of available data

The performed Domain Adaptation procedure allowed to adapt the develop ANN to real data using a limited amount of data. Indeed, for the training procedure were used only 1000 data points, which represent the 40% of the available dataset. On the contrary, for ANN development using BIOWIN data were used 6912 measurements, which represent the 80% of the simulated dataset (total BIOWIN data points were 8640). The calculation of statistical parameters to evaluate the predictive performance of the ANN model when feed with real data is reported in Table 32,

Table 34 and Table 35 for BOD, COD, TSS, respectively. Since the dataset was limited, the training, validation and test phases were repeated 10 times (10 iteration) to evaluate the stability of the produced outputs.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Table 33: Statistical parameters calculated to evaluate the performance of the developed ANN when fed with real data after domain adaptation procedure for BOD data prediction for the test phase

Fold	CC	RMSE	SI	BIAS	MeanAPD	StdAPD
1	0.333	2.081	0.350	204.791	25.955	23.617
2	0.178	1.944	0.345	510.134	26.650	17.764
3	0.102	2.712	0.444	-1.077	25.920	25.960
4	0.458	2.994	0.423	-667.871	30.623	34.883
5	0.596	2.091	0.350	217.192	28.693	42.896
6	0.078	2.724	0.438	-60.828	26.697	25.589
7	0.342	3.032	0.437	-683.296	27.777	19.246
8	0.232	2.197	0.371	125.240	25.318	21.162
9	-0.053	2.545	0.434	115.646	24.058	24.525
10	0.350	2.195	0.363	27.128	26.505	30.610
Mean	**0.262**	**2.452**	**0.395**	**-21.294**	**26.820**	**26.625**

Table 34: Statistical parameters calculated to evaluate the performance of the developed ANN when fed with real data after domain adaptation procedure for COD data prediction for the test phase

Fold	CC	RMSE	SI	BIAS	MeanAPD	StdAPD
1	0.137	4.960	0.289	354.772	23.287	19.715
2	0.127	5.705	0.337	1370.614	27.692	23.672
3	0.064	5.256	0.291	-335.619	22.294	18.237
4	0.160	5.229	0.285	-1538.485	19.268	15.028
5	0.092	6.141	0.335	-540.931	24.026	21.141
6	0.284	5.175	0.300	915.706	26.094	23.744
7	-0.076	5.995	0.343	211.393	23.150	19.481
8	0.120	4.999	0.292	339.115	21.086	18.871
9	0.098	3.872	0.230	-494.672	17.206	13.108
10	0.198	4.757	0.273	-311.962	22.135	17.488
Mean	**0.120**	**5.209**	**0.298**	**-3.007**	**22.624**	**19.048**

Table 35: Statistical parameters calculated to evaluate the performance of the developed ANN when fed with real data after domain adaptation procedure for TSS data prediction for the test phase

Fold	CC	RMSE	SI	BIAS	MeanAPD	StdAPD

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

1	0.331	1.550	0.280	19.210	21.311	23.049
2	-0.016	1.721	0.342	374.585	24.037	24.443
3	0.325	1.580	0.297	-79.725	18.729	26.972
4	0.177	1.788	0.332	36.210	18.972	30.082
5	0.324	1.449	0.254	-373.082	12.649	16.762
6	0.261	2.145	0.385	-97.249	25.017	34.198
7	0.396	1.521	0.269	-34.581	19.796	25.021
8	0.317	1.058	0.199	39.520	16.163	15.767
9	0.105	1.143	0.215	11.814	15.407	12.986
10	0.166	1.478	0.267	23.510	18.572	17.740
Mean	**0.238**	**1.543**	**0.284**	**-7.979**	**19.065**	**22.702**

Unlikely, results obtained during the test phase were not consistent during the 10 different iterations, and the model was not able to produce stable outputs as indicated by the very different values of CC and error indexes (Table 32,

Table 34 and Table 35) obtained for all the three target parameters. According to the scientific literature, in the case of real sensors installed on-line, acceptable errors in terms of *APD* values are assumed to be < 20% (Cecconi et al., 2019). In the case of the developed soft sensor in this work, the calculated *MeanAPD* in the 10 iterations ranged between 17 and 27% for BOD, 24 and 30% for COD, and 15 and 25% for TSS, whereas CC values were not very high probably because affected by single extreme values. Graphical results for the best and worst prediction scenarios in terms of CC are reported in Figure 84, Figure 85, Figure 86 for BOD, COD, TSS, respectively.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

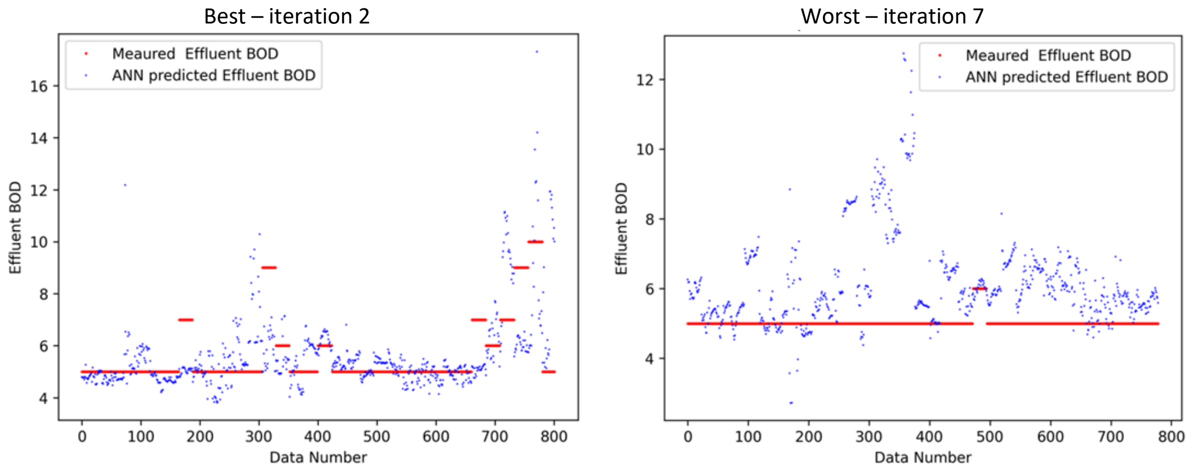


Figure 84: Comparison between BOD hourly data predicted by ANN and laboratory data for the best and worst prediction in terms of correlation coefficient

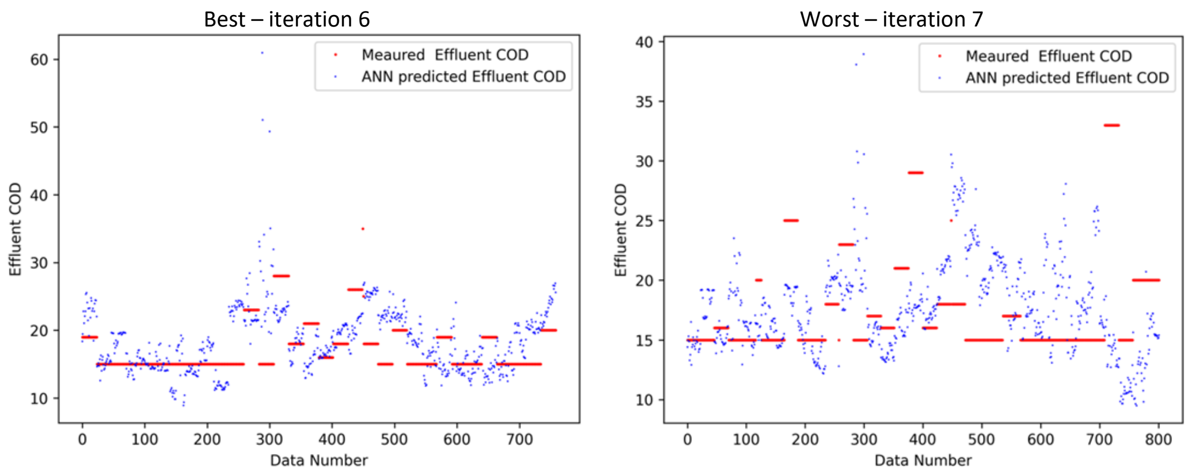


Figure 85: Comparison between COD hourly data predicted by ANN and laboratory data for the best and worst prediction in terms of correlation coefficient

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

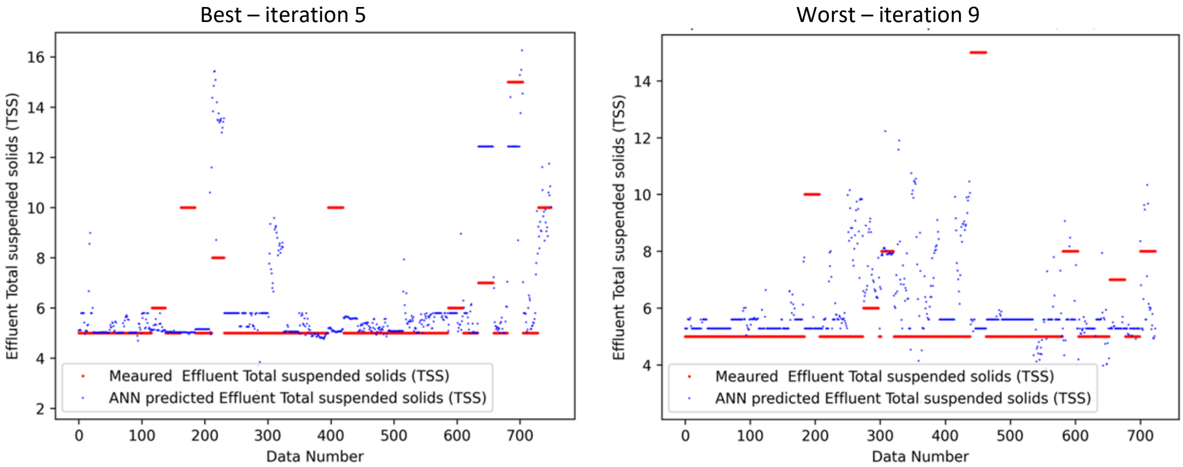


Figure 86: Comparison between TSS hourly data predicted by ANN and laboratory data for the best and worst prediction in terms of correlation coefficient

Reason for the non-very accurate results obtained after the domain adaptation procedure is the limited dataset of available real data, which were also related to non-consecutive days. In addition, it is preferable to have data measured by sensors in a continuous mode instead of daily averaged laboratory data. In this case, the soft sensor may be continuously feed with sensor data, and they may be used to support real sensors during maintenance period or fault measurements.

3.3.2.3. Development of soft sensor for the forecasting of TSS concentration

Aim of the development of a second soft sensor, in this work, was the prediction of TSS concentration in the final effluent of Peschiera Borromeo WWTP earlier in time, and specifically, 1 h, 3 h and 6 h before the occurrence of the real-time measurement. Indeed, the earlier prediction may give technicians the possibility to stop water reuse or take preventive actions before the occurrence of a hazardous event for water reuse (e.g., high TSS concentration that reduces the UV disinfection performance). For the development of this soft-sensor only simulated data by BIOWIN were used since reliable and consecutive real data by sensors were not available. Hence, an artificial neural network model for timeseries forecasting was developed using 70% of BIOWIN data for training, 10% for validation and 20% for testing. Data were split according to time of occurrence. Hence, 70% of data used for training were related to measurements from January 2020 to September 2020, measurements related to September-October 2020 were used for validation and measurements from October to December were used for testing. The developed ANN was feed with data related to Influent flowrate, Influent pH, temperature in the anoxic BIOFOR reactor, dissolved oxygen (DO) concentration in aerated BIOFOR reactor, effluent flowrate, effluent pH, and effluent TSS. The target parameter was TSS concentration in the effluent 6h, 3h and 1h before the real measurement.

The utilized Deep Learning model was the recurrent neural network (RNN) (Yu et al., 2019). The typical feature of the RNN architecture is a cyclic connection, which enables the RNN to

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

possess the capacity to update the current state based on past states and current input data (Siami-Namini et al., 2019). It means that predicted data will be calculated with respect to the past and current data. The selected RNN was a Long Short-Term Memory (LSTM) network with an internal layer with 32 cells.

A windowing procedure with the following features (Figure 87) was applied to train the model:

- Input width = 24 (it is the number of input data utilized to train the model, in this case 24 data for 24 consecutive hours);
- Label width = 1 (it is the time step for prediction);
- Offset = 1 (for 1 h prediction), 3 (for 3 h prediction), 6 (for 6 h prediction) represents the time offset of the prediction.

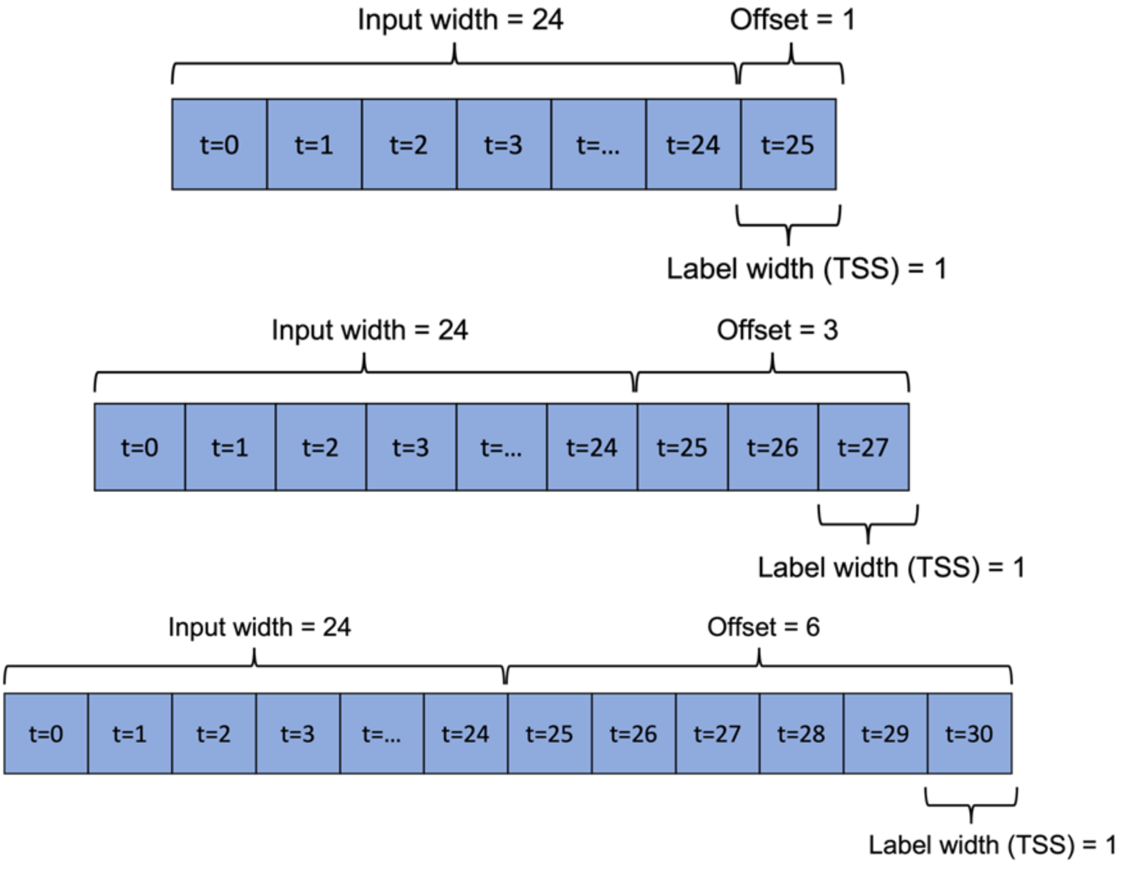


Figure 87: Features of the windowing procedure utilized to train the predictive ANN model

Statistical parameters (i.e., RMSE, SI, BIAS and MeanAPD) were calculated to evaluate the predictive performance of the ANN model along the whole periods of data. Averages of performance parameters for all utilized 24 hours windows are reported in Table 36 for the predictive model at 1h, 3h and 6h. The parameters CC was not calculated since the presence of continues constant values in some simulated periods produced indefinite values.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Table 36: Average of statistical parameters calculated to evaluate the performance of the forecasting models for TSS prediction

Forecasting model	RMSE	SI	BIAS	MeanAPD	StdAPD
1 h	0.224	0.025	-0.087	1.682	1.763
3 h	0.229442	0.0264834	-0.045	1.792	1.844
6 h	0.234	0.027	-0.099	1.855	1.661

By the analysis of Table 36 is evident that model was able to predict TSS with very high accuracy for the three different time offsets considered in this study.

For completeness of the study, performance indicators of the 1h forecasting models are shown for three different 24 h windows in Table 37, whereas predicted and true values are compared in Figure 88.

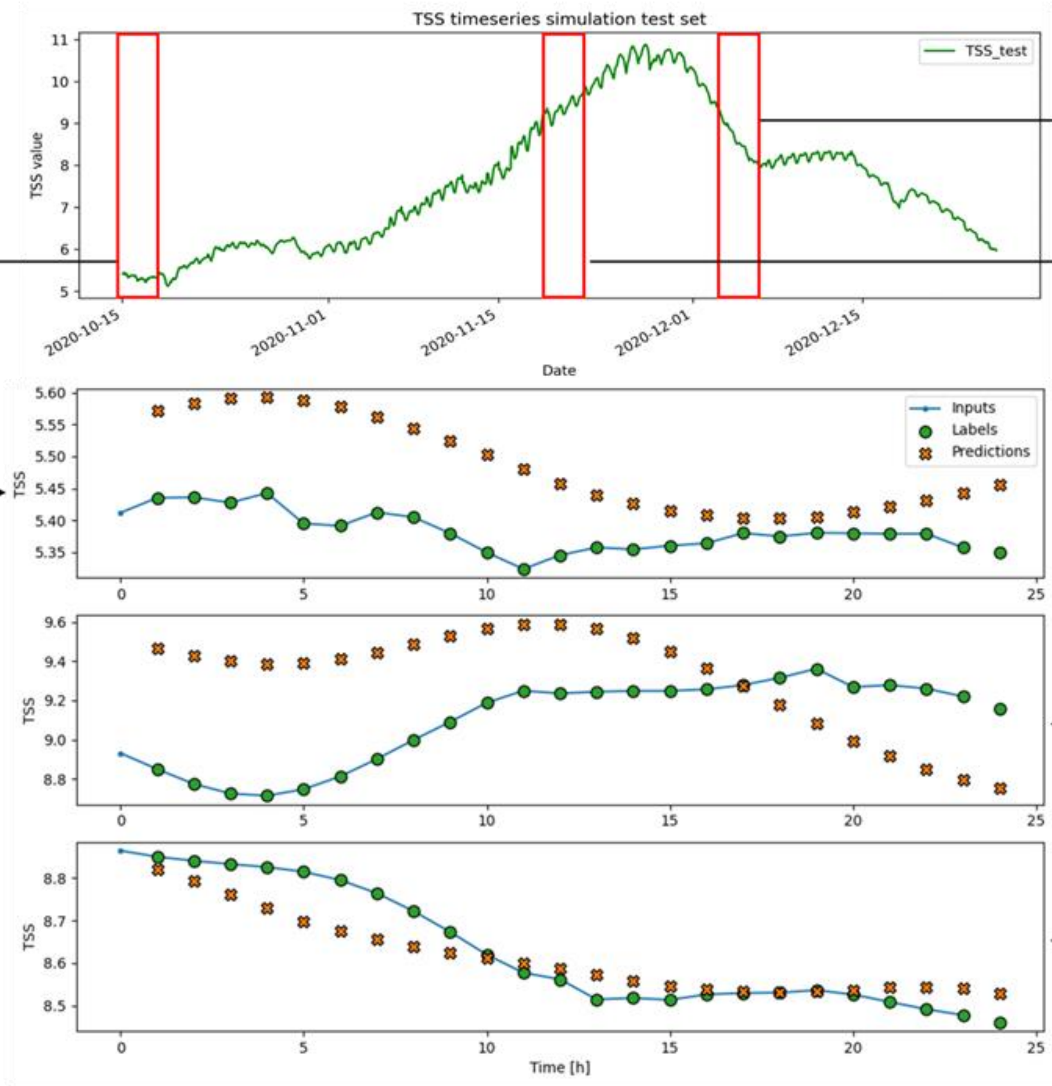


Figure 88: Comparison between predicted and true values for TSS in the case of forecasting with 1 h of time offset

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Table 37: Calculated performance indicators for the 1h forecasting model for three different 24 h windows

Window	RMSE	SI	BIAS	MeanAPD	StdAPD
1	0.117	0.022	0.103	1.924	1.011
2	0.439	0.048	0.208	4.437	2.115
3	0.060	0.007	-0.013	0.554	0.408

Similar information is shown for the 3 h forecasting model (Figure 89 and Table 38). and the 6 h forecasting model (Figure 90 and Table 39).

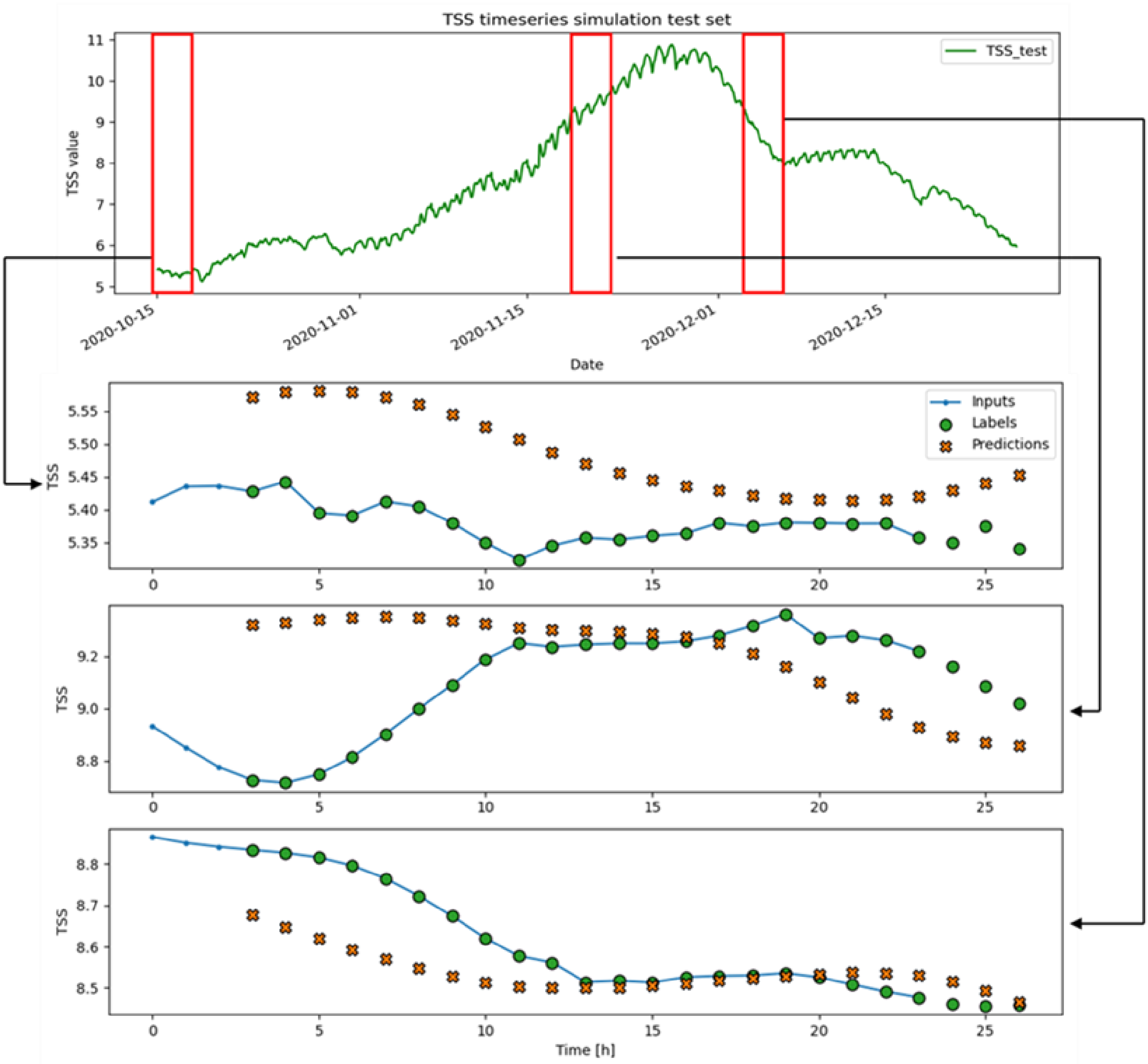


Figure 89: Comparison between predicted and true values for TSS in the case of forecasting with 3 h of time offset

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Table 38: Calculated performance indicators for the 3h forecasting model for three different 24 h windows

Window	RMSE	SI	BIAS	MeanAPD	StdAPD
1	0.120	0.022	0.107	1.998	1.000
2	0.305	0.033	0.076	2.671	2.163
3	0.103	0.012	-0.055	0.861	0.805

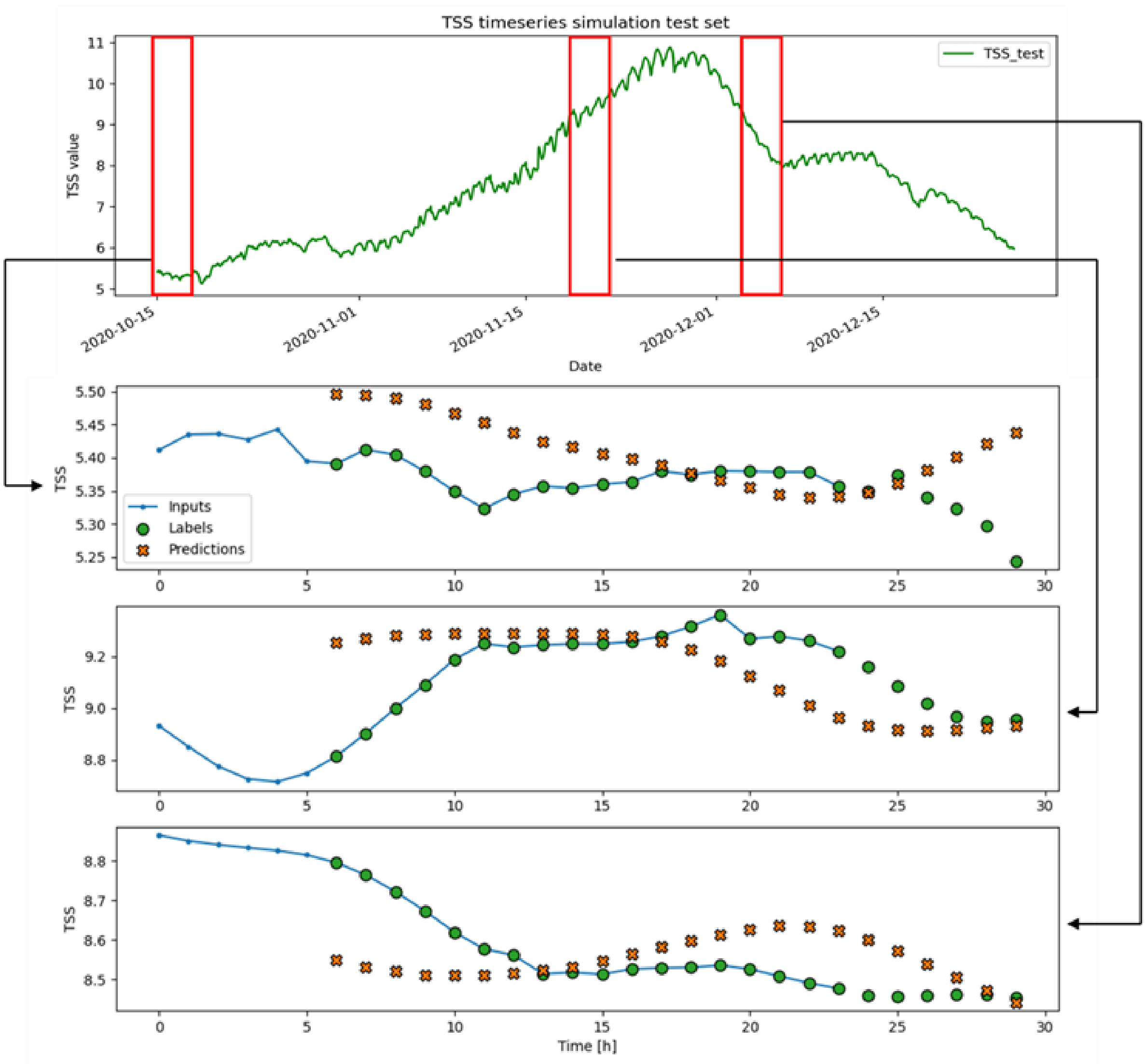


Figure 90: Comparison between predicted and true values for TSS in the case of forecasting with 6 h of time offset

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Table 39: Calculated performance indicators for the 6h forecasting model for three different 24 h windows

Window	RMSE	SI	BIAS	MeanAPD	StdAPD
1	0.079	0.014	0.051	1.182	0.911
2	0.181	0.019	-0.005	1.543	1.294
3	0.116	0.013	0.005	1.105	0.775

3.3.3. Early Warning System integration in FIWARE

Real-time sensor data from WWTP will be ingested by a dedicated back-end deployed as a serverless application. The ingestion of data triggers the processing using a dedicate event bus. As shown in Figure 91 a dedicated function written in python can run as a function in the cloud. The lambda function uses the images stored in an Elastic Container Registry (ECR) and it makes easier to update the code / models. The output of ANN is a timeseries that includes forecast to be used by the EWS. The EWS can then provide warnings depending on the predicted outputs.

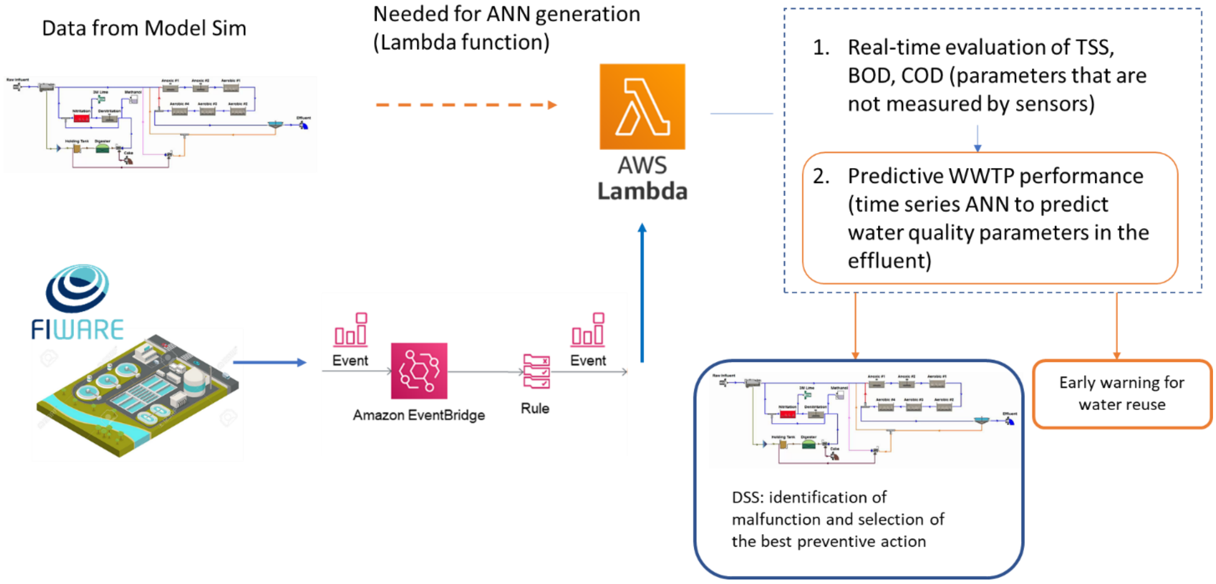


Figure 91: EWS architecture

The ANN can produce for the WebGIS information about predicted effluent wastewater quality. The warning provided by the EWS can be communicated to the WebGIS, which will indicate if the plant is able to provide reclaimed wastewater for reuse applications.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

4. Conclusions

In this report is described the implementation of Early Warning Systems designed to assure health protection in water reuse practices and during recreational activities in bathing water sites.

In the city of Paris, the EWS aims to reduce the health risks related to microbiological contamination in bathing water sites from combined sewer overflows and/or other wastewater discharges. The designed EWS is able to predict the bathing water quality in terms of E. coli concentration. Particularly, two different models have been implemented for the scope. A deterministic model (ProSe) is able to produce input data for a statistical tool using quantitative and qualitative data from upstream tributaries, WWTP discharges, combined sewer overflows (CSO), etc. On the other side, the statistical model is based on machine learning and Bayesian regression algorithms, and it is able to estimate E. Coli concentration in the selected bathing site defining the sanitary quality of water according to the Bathing Water Directive. Predicted data can be sent to dedicated app to inform managers, stakeholders and citizens about the need to close the access to the bathing site due to the occurrence of potential risks for swimmers' health.

The EWS for safe water reuse is a tool conceived within the risk-based management framework of sanitation systems. It aims at preventing bacterial and toxic contamination linked to the reuse of treated wastewater for agricultural irrigation based on: (i) a comprehensive network of multi-parameter sensors at a WWTP, (ii) new sensors for real-time and in-situ measurements (e.g., E. coli measurement - Digital Solution DS1); (iii) machine learning and forecasting algorithms to predict contamination events. Particularly, the Early Warning System designed for the case study of Milan at Peschiera Borromeo WWTP is able to provide warnings and alarms if estimated concentrations for target water quality parameters do not assure compliance with water reuse standards or are related to hazard events that may determine non- tolerable risks for human health during agricultural irrigation.

To effectively design the EWS in Milan, a semi-quantitative and a quantitative risk assessment were accomplished to define relevant hazards or the occurrence of hazardous event that may entail non-tolerable risks for human health. Outcomes of the semi quantitative risk assessment highlighted the effectiveness of control measures to minimize health risk and proposed additional measures that may be further applied to reduce still relevant risks, such as the presence of hazardous pathogens in wastewater due to failure of the disinfection system. A Failure Modes-Effects Analysis (FMEA) excluded the presence of possible chemical hazards in the treated wastewater. On the contrary, outcomes of the performed quantitative microbial risk analysis (QMRA), which was performed using the Monte Carlo approach, highlighted the occurrence of a non-negligible risk for rotavirus if appropriate log-removals for virus deactivation cannot be assured at the WWTP. In this context, the EWS system represents a strategical tool to guarantee a continuous supply of safe treated water. The EWS implemented in Milan can generate warning and alarms related to measurements obtained by on-line sensors and from machine learning algorithms (i.e., soft sensors). Particularly, the developed soft sensors are able to predict the following parameters: (i) Biochemical oxygen

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

demand (BOD – real-time prediction); (ii) Chemical oxygen demand (COD – real-time prediction); (iii) Total suspended solids (TSS – real-time prediction and up to 6 hours earlier prediction). To predict these target parameters, the elaborated soft-sensors use data on Influent flowrate, Influent pH, temperature in the anoxic reactor, dissolved oxygen (DO) concentration in aerated BIOFOR reactor, effluent flowrate, effluent pH, which are measured by on-line probes at Peschiera Borromeo WWTP.

The development of the two solutions was contextualised in both cases into a risk management approach, also considering the legislative requirements of the European Bathing Water Directive and Water Reuse Regulation. Overall, the risk management approaches promote a process control to complement end-product quality testing by avoiding the final provision of not suitable products. Hence, the risk management process can be defined as the identification, evaluation, and prioritization of risks, followed by the coordinated application of actions/resources to minimize, monitor, and control the probability or impact of unfortunate events. In this context, an EWS can be considered as an auxiliary tool for risk management, which allows the identification of the occurrence of a contamination/hazard event in real-time or even beforehand contributing in this way to minimize risks. Hence, the development of an EWS is strictly connected with the realization of a risk management plan, where care has to be taken in the phases of system description and hazard identification. To reach this scope the expertise of a multidisciplinary team that include engineers, chemists, experts in environmental science, biology, agriculture and technical operators of the water sector results strategical. Particularly, parameters to monitor by EWS have to be representatives of hazards that may occur in the described system, and they have to be defined according to the outcomes of the risk analysis or based on legislative requirements. Indeed, the advantage to implement an EWS based on a risk assessment to monitor water reuse or bathing water quality relies in the possibilities to predispose a well-designed monitoring (in some cases even predictive) system able to control all the possible hazardous events that may occur in a specific context. Furthermore, an EWS can also absolve the function of decision support system providing specific and defined information/instructions to minimize the occurring risk.

Main difficulties that have been faced to develop EWS are related to phase of system description and hazard identification. Particularly, in the case of water reuse very few information was available about the list of possible malfunctions or contamination events that may occur during the operation of a WWTP and the related frequency of occurrence. A register for the annotation of malfunctions and contamination events is paramount for the realization of a risk management plan, and consequently for the implementation of an effective EWS.

In DWC, the development of a EWS was based on the collection of real-time data by sensors and on the use of machine-learning algorithms, which represent very powerful tools to complement and improve the information obtained by probe measurements. However, the development of accurate mathematical models for monitoring and prediction is based on the availability of robust and reliable dataset, which are rarely available in real contexts. For example, scarce maintenance of probes can impact data reliability, and it results to be a very frequent and significant issue. Furthermore, very few data are available about the occurrence

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

of malfunctions or contamination events, and this fact can limit the modeling of these hazardous events. The utilization of mechanistic and physical model can in part compensate for this lack. The roadmap for the replicability of the proposed solutions (and hence of the potential standardization of the risk-based approach) relies on the definition of specific guidelines for risk management plan preparation. Particularly, defined procedures for the annotation of contamination events or malfunctions occurrences may be set up and provided to interested stakeholders to help them in the identification and recording of the occurrence of hazardous events. In addition, specific protocols for the validation of real-time probe measurements should be defined. Indeed, the availability of good quality data is the first and unavoidable requirement for the replicability of the proposed solutions in other locations.

In this work the risk assessment for water reuse and consequently the implementation of the EWS has been limited to the operation of the water reclamation facility. However, for an exhaustive evaluation of risks, the assessment should be extended to other components of the water reuse systems, including pipes and storage tanks that may be needed to distribute treated wastewater to users, and to critical points of the sewer network that convey the raw wastewater to the WWTP. Hence, an EWS may be implemented to monitor those other components of the reuse system.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

References

A.H. Havelaar, J.M.Melse. (2003). Quantifying Public Health Risk in the WHO Guidelines for Drinking Water Quality. *WHO RIVM Report*, *June 2014*, 1–49.

Allard, A., & Vantarakis, A. (2018). Adenoviruses. In: J.B. Rose and B. Jiménez-Cisneros, (eds) Water and Sanitation for the 21st Century: Health and Microbiological Aspects of Excreta and Wastewater Management. (J.S Meschke, and R. Girones (eds), Part 3: Specific Excreted Pathogens: Environ. *Global Water Pathogen Project, Michigan State University, E. Lansing, MI, UNESCO*, 3–10. https://www.waterpathogens.org/book/ascaris

Anastasi, E. M., Wohlsen, T. D., Stratton, H. M., & Katouli, M. (2013). Survival of Escherichia coli in two sewage treatment plants using UV irradiation and chlorination for disinfection. *Water Research*, *47*(17), 6670–6679. https://doi.org/10.1016/j.watres.2013.09.008

Bahri, A., Drechsel, P., Raschid-Sally, L., & Redwood, M. (2010). Wastewater Irrigation and Health. Assessing and Mitigating Risk in Low-income Countries. *Routledge*, 432.

Bailey, E. S., Casanova, L. M., Simmons, O. D., & Sobsey, M. D. (2018). Tertiary treatment and dual disinfection to improve microbial quality of reclaimed water for potable and non-potable reuse: A case study of facilities in North Carolina. *Science of The Total Environment*, *630*, 379–388. https://doi.org/10.1016/j.scitotenv.2018.02.239

Bedoui, K. (2014). *Good Water Governance in Wastewater Reuse in case study: Medenine, Tunisia. In Sustainable Integrated Wastewater Treatment and Reuse. Sharm El-Shaikh*.

Bergmeir, C., & Benítez, J. M. (2012). On the use of cross-validation for time series predictor evaluation. *Information Sciences*, *191*, 192–213. https://doi.org/10.1016/j.ins.2011.12.028

Betancourt, W. Q., & Shulman, L. M. (2017). Polioviruses and other Enteroviruses. In: J.B. Rose and B. Jiménez-Cisneros, (eds) Water and Sanitation for the 21st Century: Health and Microbiological Aspects of Excreta and Wastewater Management. (J.S Meschke, and R. Girones (eds), Part 3: Specific Exc. *Global Water Pathogen Project, Michigan State University, E. Lansing, MI, UNESCO*.

Black, G., Jones, M., Vale, P., Johnson, N., Nocker, A., Cartmell, E., & Dotro, G. (2014). Biofilm Responses to Toxic Shocks in Closed Pipes: Using Nitrous Oxide Emissions as an Early Warning of Toxicity Ahead of a Wastewater Treatment Works. *Water, Air, & Soil Pollution Soil Pollut*, *225*, 1837.

Bonadonna, L., Briancesco, R., Ottaviani, M., & Veschetti, E. (2002). Occurrence of Cryptosporidium Oocysts in Sewage Effluents and Correlation with Microbial, Chemical and Physical Water Variables. *Environmental Monitoring and Assessment*, *75*, 241–252.

Bonetta, S., Pignata, C., Gasparro, E., Richiardi, L., Bonetta, S., & Carraro, E. (2022). Impact of wastewater treatment plants on microbiological contamination for evaluating the risks of wastewater reuse. *Environmental Sciences Europe*, *34*(1). https://doi.org/10.1186/s12302-022-00597-0

Briancesco, R., & Bonadonna, L. (2005a). An Italian study on Cryptosporidium and Giardia in wastewater, fresh water and treated water. *Environmental Monitoring and Assessment*, *104*(1–3), 445–457. https://doi.org/10.1007/s10661-005-2282-4

Briancesco, R., & Bonadonna, L. (2005b). An Italian Study on Cryptosporidium and Giardia in Wastewater, Fresh Water and Treated Water. *Environmental Monitoring and Assessment*, *104*(1–3), 445–457. https://doi.org/10.1007/s10661-005-2282-4

Brussen, M. (2007). On-line water quality monitoring. *Review of Sydney's Current Status and Future Needs Sydney Water Report.(Sydney)*.

Carraro, E., Fea, E., Salva, S., & Gilli, G. (2000). Impact of a wastewater treatment plant on Cryptosporidium oocysts and Giardia cysts occurring in a surface water. *Water Science and Technology*, *41*(7), 31–37. https://doi.org/10.2166/wst.2000.0112

Cecconi, F., Reifsnyder, S., Ito, Y., Jimenez, M., Sobhani, R., & Rosso, D. (2019). ISE-ammonium sensors in WRRFs: field assessment of their influencing factors. *Environ. Sci.: Water Res. Technol.*, *5*, 737–746.

Chahal, C., van den Akker, B., Young, F., Franco, C., Blackbeard, J., & Monis, P. (2016). *Pathogen and Particle Associations in Wastewater* (pp. 63–119). https://doi.org/10.1016/bs.aambs.2016.08.001

Chow, C., Liu, J., Li, J., Swain, N., & Reid, K. (2018). Development of smart data analytics tools to support wastewater treatment plant operation. *Chemometrics and Intelligent Laboratory Systems*, *177*, 140–150.

Corpuz, M. V. A., Buonerba, A., Vigliotta, G., Zarra, T., Ballesteros, F., Campiglia, P., Belgiorno, V., Korshin, G., & Naddeo, V. (2020). Viruses in wastewater: occurrence, abundance and detection methods. *Science of the Total Environment*, *745*, 140910. https://doi.org/10.1016/j.scitotenv.2020.140910

De Souza, J. B., Queiroz Valdez, F., Jeranoski, R. F., De Sousa Vidal, C. M., & Cavallini, G. S. (2015). Water and Wastewater Disinfection with Peracetic Acid and UV Radiation and Using Advanced Oxidative Process PAA/UV. *International Journal of Photoenergy*, *2015*. https://doi.org/10.1155/2015/860845

DEMOWARE. (2013). *Innovation Demonstration for a Competitive and Innovative European Water Reuse Sector*. *1*(619040).

DEMOWARE project. (2013). *D 3 . 3 : Generic assessment of treat- ment trains concerning their envi- ronmental impact and risk reduction potential*. *1*(619040).

Deng, S., Yan, X., Zhu, Q., & Liao, C. (2019). The utilization of reclaimed water: Possible risks arising from waterborne contaminants. *Environmental Pollution*, *254*, 113020. https://doi.org/10.1016/j.envpol.2019.113020

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Du, Y., Chen, Y., Zou, L., Deng, S., Li, G., & Zang, D. (2019). Monitoring the Activated Sludge Activities Affected by Industrial Toxins via an Early-Warning System Based on the Relative Oxygen Uptake Rate (ROUR) Index. *Applied Sciences*, *9*, 154.

Bathing Water Directive, (2006).

EU Regulation. (2020). *REGULATION (EU) 2020/741 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 25 May 2020 on minimum requirements for water reuse* (Vol. 2019, Issue April).

Farahani, A., Voghoei, S., Rasheed, K., & Arabnia, H. R. (2021). *A Brief Review of Domain Adaptation*. 877–894. https://doi.org/10.1007/978-3-030-71704-9_65

Francy, D. S., Stelzer, E. A., Bushon, R. N., Brady, A. M. G., Williston, A. G., Riddell, K. R., Borchardt, M. A., Spencer, S. K., & Gellner, T. M. (2012). Comparative effectiveness of membrane bioreactors, conventional secondary treatment, and chlorine and UV disinfection to remove microorganisms from municipal wastewaters. *Water Research*, *46*(13), 4164–4178. https://doi.org/10.1016/j.watres.2012.04.044

German Water Partnership. (n.d.). *WATER 4.0*.

Goodwin, D., Raffin, M., Jeffrey, P., & Smith, H. M. (2015). Applying the water safety plan to water reuse: Towards a conceptual risk management framework. *Environmental Science: Water Research and Technology*, *1*(5), 709–722. https://doi.org/10.1039/c5ew00070j

Grayman, W. M., Deininger, R. A., & Males, R. M. (2001). *Design of early warning and predictive source-water monitoring systems*. American Water Works Association.

Gruppo di lavoro ISS Ambiente Rifiuti. (2020). Indicazioni ad interim su acqua e servizi igienici in relazione alla diffusione del virus SARS-CoV-2. *Novum Testamentum Graece*, *Rapporti I*. https://doi.org/10.1017/cbo9781139854634.009

Hamilton, K. A., Waso, M., Reyneke, B., Saeidi, N., Levine, A., Lalancette, C., Besner, M., Khan, W., & Ahmed, W. (2018). *Cryptosporidium* and *Giardia* in Wastewater and Surface Water Environments. *Journal of Environmental Quality*, *47*(5), 1006–1023. https://doi.org/10.2134/jeq2018.04.0132

Hassaballah, A. H., T, B., Nyitrai, J., Dai, N., & Sassoubre, L. (2019). Inactivation of E. coli, Enterococcus spp., somatic coliphage, and Cryptosporidium parvum in wastewater by peracetic acid (PAA), sodium hypochlorite, and combined PAA-ultraviolet disinfection. *Water Research & Technology*.

Hijnen, W. A. M., Beerendonk, E. F., & Medema, G. J. (2006). Inactivation credit of UV radiation for viruses, bacteria and protozoan (oo)cysts in water: A review. *Water Research*, *40*(1), 3–22. https://doi.org/10.1016/j.watres.2005.10.030

Hong, S., Choi, I., Lim, B., & Kim, H. (2012). A DO- and pH-Based Early Warning System of Nitrification inhibition for biological nitrogen removal processes. *Sensors*, *12*, 16334–16352.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Howard, I., Espigares, E., Lardelli, P., Martín, J. L., & Espigares, M. (2004). Evaluation of microbiological and physicochemical indicators for wastewater treatment. *Environmental Toxicology*, *19*(3), 241–249. https://doi.org/10.1002/tox.20016

ISO 10705-2. (2000). *Water quality — Detection and enumeration of bacteriophages — Part 2: Enumeration of somatic coliphages*.

ISO 15553. (2006). *Water quality — Isolation and identification of Cryptosporidium oocysts and Giardia cysts from water*.

ISO 16075-1. (2020). *Guidelines for treated wastewater use for irrigation projects — Part 1: The basis of a reuse project for irrigation*.

ISO 17995. (2019). *Water quality — Detection and enumeration of thermotolerant Campylobacter spp*.

ISO 19250. (2010). *Water quality — Detection of Salmonella spp.*

ISO 9308-2. (2012). *Water quality — Enumeration of Escherichia coli and coliform bacteria — Part 2: Most probable number method*.

Jothikumar, N., Lowther, J. A., Henshilwood, K., Lees, D. N., Hill, V. R., & Vinjé, J. (2005). Rapid and sensitive detection of noroviruses by using TaqMan-based one-step reverse transcription-PCR assays and application to naturally contaminated shellfish samples. *Applied and Environmental Microbiology*, *71*(4), 1870–1875. https://doi.org/10.1128/AEM.71.4.1870-1875.2005

Jurga, A., Gemza, N., & Janiak, K. (2017). A concept development of an early warning. *E3S Web of Conferences*.

Katayama, H., & Vinjé, J. (2018). Norovirus and other Calicivirus. In: J.B. Rose and B. Jiménez-Cisneros, (eds) Water and Sanitation for the 21st Century: Health and Microbiological Aspects of Excreta and Wastewater Management . (J.S Meschke, and R. Girones (eds), Part 3: Specific Excrete. *Global Water Pathogen Project, Michigan State University, E. Lansing, MI, UNESCO.*, 3–10. https://www.waterpathogens.org/book/ascaris

Kay, D., Crowther, J., Stapleton, C. M., Wyer, M. D., Fewtrell, L., Edwards, A., Francis, C. A., McDonald, A. T., Watkins, J., & Wilkinson, J. (2008). Faecal indicator organism concentrations in sewage and treated effluents. *Water Research*, *42*(1–2), 442–454. https://doi.org/10.1016/j.watres.2007.07.036

Landi, S., & Baroncelli, P. (2000). *L ’ acqua irrigua : campionamento , analisi chimico-fisiche e interpretazione dei risultati L ’ analisi può essere fatta in qualsiasi periodo I parametri analitici rilevanti*.

Langergraber, G., Alex, J., Weissenbacher, N., Woerner, D., Ahnert, M., Frehmann, T., Halft, N., Hobus, L., Plattes, M., Spering, V., & Winkler, S. (2008). Generation of diurnal variation for influent data for dynamic simulation. *Water Science and Technology*, *57*(9), 1483–1486. https://doi.org/10.2166/wst.2008.228

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Lemarchand, K., & Lebaron, P. (2003). Occurrence of *Salmonella* spp. and *Cryptosporidium* spp. in a French coastal watershed: relationship with fecal indicators. *FEMS Microbiology Letters*, *218*(1), 203–209. https://doi.org/10.1111/j.1574-6968.2003.tb11519.x

Liberti, L., Notarnicola, M., & Petruzzelli, D. (2003). Advanced treatment for municipal wastewater reuse in agriculture. UV disinfection: parasite removal and by-product formation. *Desalination*, *152*(1–3), 315–324. https://doi.org/10.1016/S0011-9164(02)01079-2

Lin, J., & Ganesh, A. (2013). Water quality indicators: Bacteria, coliphages, enteric viruses. *International Journal of Environmental Health Research*, *23*(6), 484–506. https://doi.org/10.1080/09603123.2013.769201

Lu, X., & Erdman, D. (2006). Molecular typing of human adenoviruses by PCR and sequencing of a partial region of the hexon gene. *Archives of Virology*, *151*, 1587–1602.

Malorgio, F. (2004). Uso razionale delle risorse nel florovivaismo: l'acqua. In *Quaderno ARSIA*. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Uso+razionale+delle+risorse+nel+florovivaismo:+l'acqua+•#0

Mara, D. (2006). *Guidance note for Programme Managers and Engineers A NumericAl Guide to Volume 2 of the GuideliNes ANd PrActicAl AdVice oN how to trANsPose them iNto NAtioNAl*. *2*(January 2008), 1–12.

Marín, I., Goñi, P., Lasheras, A. M., & Ormad, M. P. (2015). Efficiency of a Spanish wastewater treatment plant for removal potentially pathogens: Characterization of bacteria and protozoa along water and sludge treatment lines. *Ecological Engineering*, *74*, 28–32. https://doi.org/10.1016/j.ecoleng.2014.09.027

Metcalf & Eddy. (2014). *Wastewater Engineering: Treatment and Resource Recovery* (N. Y. McGraw-Hill (ed.); 5th Editio).

NRMMC-EPHC-AHMC. (2006). *Australian Guidelines for Water Recycling: Managing Health and Environmental Risks (Phase1)*.

Oakley, S., & Mihelcic, J. (2019). Pathogen Reduction and Survival in Complete Treatment Works. In *Global Water Pathogen Project*. https://doi.org/10.14321/waterpathogens.49

Pina, S., Puig, M., Lucena, F., Jofre, J., & Girones, R. (1998). Viral pollution in the environment and in shellfish: Human adenovirus detection by PCR as an index of human viruses. *Applied and Environmental Microbiology*, *64*(9), 3376–3382. https://doi.org/10.1128/aem.64.9.3376-3382.1998

Poulin, M., Servais, P., Mouchel, J., Thérial, C., & Lesage, L. (2013). *Modélisation de la contamination fécale en Seine : impact des rejets de temps de pluie*. *February*.

Qiu, Y., Li, Q., Lee, B. E., Ruecker, N. J., Neumann, N. F., Ashbolt, N. J., & Pang, X. (2018). UV inactivation of human infectious viruses at two full-scale wastewater treatment plants in Canada. *Water Research*, *147*, 73–81. https://doi.org/10.1016/j.watres.2018.09.057

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

R3WATER. (2017a). *Water in the circular economy – innovations for urban water treatment-Reearch project founded by the European Union*.

R3WATER. (2017b). *Water in the circular economy – innovations for urban water treatment*.

Ramo, A., Del Cacho, E., Sánchez-Acedo, C., & Quílez, J. (2017). Occurrence and genetic diversity of Cryptosporidium and Giardia in urban wastewater treatment plants in north-eastern Spain. *Science of The Total Environment*, *598*, 628–638. https://doi.org/10.1016/j.scitotenv.2017.04.097

Rebelo, A., Farabegoli, G., Andreotti, F., Balmer, J., Vella, M., Ronald Van Tunen, R., Stuart Gunput, S., Stella Perikenti, S., & Ece, P. (2018). *Report on Urban Water Reuse*.

Ruder, S. (2017). *An Overview of Multi-Task Learning in Deep Neural Networks*. *May*. http://arxiv.org/abs/1706.05098

Sakson, G., Brzezinska, A., & Kowalski, K. (2019). Monitoring, early warning and sustainable management system for lodz wastewater treatment plant as a water protection tool. *JCEEA*.

Santiago, P., Jiménez-Belenguer, A., García-Hernández, J., Estellés, R. M., Hernández Pérez, M., Castillo López, M. A., Ferrús, M. A., & Moreno, Y. (2018). High prevalence of Salmonella spp. in wastewater reused for irrigation assessed by molecular methods. *International Journal of Hygiene and Environmental Health*, *221*(1), 95–101. https://doi.org/10.1016/j.ijheh.2017.10.007

Seis, W, Zamzow, M., Caradot, N., & Rouault, P. (2018). On the implementation of reliable early warning systems at European bathing water using multivariate Bayesian regression modelling. *Water Research*, *143*(15), 301–312.

Seis, Wolfgang. (2012). *Risk assessment of Braunschweig wastewater reuse scheme*.

Seis, Wolfgang. (2022). *QMRA (Version 0.1.3) [Computer software]* (p. https://doi.org/10.5281/zenodo.6457511).

Seis, Wolfgang, & Remy, C. (2013). *D3 . 1 Appropriate and user friendly methodologies for Risk assess- ment , Life Cycle Assessment , and Water Footprinting*. *1*(619040). http://demoware.ctm.com.es/en/results/deliverables

Servais, P., Poulin, M., Flipo, N., Lesage, L., Rocher, V., Masnada, S., Lucas, F., & Mouchel, J. M. (2010). *Vers un modèle de la contamination fécale en Seine par temps de pluie*. 1–19.

Siami-Namini, S., Tavakoli, N., & Namin, A. (2019). The Performance of LSTM and BiLSTM in Forecasting Time Series. *2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA*, 3285–3292.

Soller, J. A., Eftim, S. E., Warren, I., & Nappier, S. P. (2017). Evaluation of microbiological risks associated with direct potable reuse. *Microbial Risk Analysis*, *5*, 3–14. https://doi.org/10.1016/j.mran.2016.08.003

Souza, F. A. A., Araújo, R., & Mendes, J. (2016). Review of soft sensor methods for regression

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

applications. *Chemometrics and Intelligent Laboratory Systems*, *152*, 69–79. https://doi.org/10.1016/j.chemolab.2015.12.011

Thwaites, B. J., Short, M. D., Stuetz, R. M., Reeve, P. J., Alvarez Gaitan, J.-P., Dinesh, N., & van den Akker, B. (2018). Comparing the performance of aerobic granular sludge versus conventional activated sludge for microbial log removal and effluent quality: Implications for water reuse. *Water Research*, *145*, 442–452. https://doi.org/10.1016/j.watres.2018.08.038

Truchado, P., Garre, A., Gil, M. I., Simón-Andreu, P. J., Sánchez, G., & Allende, A. (2021). Monitoring of human enteric virus and coliphages throughout water reuse system of wastewater treatment plants to irrigation endpoint of leafy greens. *Science of the Total Environment*, *782*. https://doi.org/10.1016/j.scitotenv.2021.146837

UNI EN ISO 15216-2. (2019). *Microbiology of the food chain - Horizontal method for determination of hepatitis A virus and norovirus using real-time RT-PCR - Part 2: Method for detection (ISO 15216-2:2019)*.

US EPA. (2005). *Guidelines for Carcinogen Risk Assessment*.

US EPA. (2012). *Guidelines for water reuse*.

Vernile, A., Nabi, A. Q., Bonadonna, L., Briancesco, R., & Massa, S. (2009). Occurrence of giardia and cryptosporidium in Italian water supplies. *Environmental Monitoring and Assessment*, *152*(1–4), 203–207. https://doi.org/10.1007/s10661-008-0308-4

WERF. (2004). *WERF California*.

WHO. (2006). Safe use of wastewater , excreta and greywater guidelines. Volume 2: wastewater use in agriculture. *World Health*, *II*, 204. http://whqlibdoc.who.int/publications/2006/9241546832_eng.pdf

WHO. (2016a). *Quantitative Microbial Risk Assessment: Application for Water Safety Management*. http://www.who.

WHO. (2016b). Sanitation safety planning. *Notes and Queries*, 156.

Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures. *Neural Computation*, *3*(7), 1235–1270.

Zaho, S., Liu, P., Niu, Y., Chen, Z., Khan, A., Zhang, P., & Li, X. (2018). A Novel Early Warning System Based on a Sediment Microbial Fuel Cell for In Situ and Real Time Hexavalent Chromium Detection in Industrial Wastewater. *Sensors*, *18*(2), 642.

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Annex B - Python code of developed soft sensor that can be run in the cloud

Train ANN Multitask model on Biowin data

```
import os
import sys
import random
import numpy as np
import pandas as pd
from pathlib import Path
import tensorflow as tf
from tensorflow.keras.models import Model
from tensorflow.keras.layers import Input, Dense
from tensorflow.keras.optimizers import Adam, SGD
from tensorflow.keras.callbacks import EarlyStopping
from tensorflow.keras.backend import clear_session
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
import matplotlib.pyplot as plt
from scipy.stats.stats import pearsonr
from sklearn.metrics import mean_squared_error
from math import sqrt
import csv
from pickle import dump

random.seed(2)
random_state = 2

#WORKING DIRECTORY
path = Path('path/to/insert')
save_path = Path(path /'Ann_model_MLT')



#DATA LOADING
data = pd.read_excel(path / 'data/Biowin simulated data.xlsx', header=1)

#variables to consider
feat = ['Influent Flow',
'Influent pH',
'Biofor DN Temperature',
```


digital-water.city
digitalwater_eu
digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

```
'Biofor CN - Dissolved oxygen',
'Effluent Flow',
'Effluent pH']

targets = list(data.columns[-3:])

X = data[feat]
y = data[targets]


#PARAMETERS TO SET

Loss= {"BOD": 'mean_squared_error',
    "COD": 'mean_squared_error',
    "TSS": 'mean_squared_error'}

lr=0.01

Optimizer= Adam(learning_rate=lr)
#Optimizer= SGD(learning_rate=lr)


epochs=1000
bs=128


#DEFINE NETWORK ARCHITECTURE FUNCTION
clear_session()

def ANN_AP(X,T,Loss,Optimizer):

  x = Input(shape=(X.shape[1], ))
  shared = Dense(32, activation='relu')(x)
  sub1 = Dense(16, activation='relu')(shared)
  sub2 = Dense(16, activation='relu')(shared)
  sub3 = Dense(16, activation='relu')(shared)
  sub1 = Dense(8, activation='relu')(sub1)
  sub2 = Dense(8, activation='relu')(sub2)
  sub3 = Dense(8, activation='relu')(sub3)
  out1 = Dense(1, name="BOD", activation='linear')(sub1)
```


digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

```
    out2 = Dense(1, name="COD", activation='linear')(sub2)
    out3 = Dense(1, name="TSS", activation='linear')(sub3)

    model = Model(inputs=x, outputs=[out1, out2, out3])

    model.summary()

    model.compile(loss=Loss, optimizer=Optimizer, metrics=['MeanSquaredError','MeanAbsoluteError'])

    return model

#SPLIT AND SCALE DATA
x_tv, x_test, y_tv, y_test = train_test_split(X, y, test_size=0.25, shuffle= True, random_state = random_state)
x_train, x_valid, y_train, y_valid = train_test_split(x_tv, y_tv, test_size=0.20, shuffle= True, random_state = random_state)
print(x_train.shape, x_valid.shape, x_test.shape)

# Transform
ss = StandardScaler()
ss.fit(x_train)
x_train_scaled = ss.transform(x_train)
x_valid_scaled = ss.transform(x_valid)
x_test_scaled  = ss.transform(x_test)

# save the scaler
dump(ss, open(save_path / 'scaler_MLT.pkl', 'wb'))

#CALL AND TRAIN NETWORK
NeuNet = ANN_AP(X,y,Loss,Optimizer)

es = EarlyStopping(monitor='val_loss', patience=150, restore_best_weights=True)
history = NeuNet.fit(x_train_scaled, y={"BOD": y_train.iloc[:,0],"COD": y_train.iloc[:,1],"TSS": y_train.iloc[:,2]},
            validation_data = (x_valid_scaled,{"BOD": y_valid.iloc[:,0],"COD": y_valid.iloc[:,1],"TSS": y_valid.iloc[:,2]}),
            epochs=epochs, batch_size=bs, verbose=1, callbacks=[es])

Save trained model
NeuNet.save(save_path)

PLOT TRAINING CURVES
mae = history.history['mean_absolute_error']
val_mae = history.history['val_mean_absolute_error']
```


digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

```
loss = history.history['loss']
val_loss = history.history['val_loss']
lista = [loss,val_loss,mae,val_mae]
epochs = range(len(loss))

plt.figure()
plt.plot(epochs, mae, 'b', label='Training mae')
plt.plot(epochs, val_mae, 'r', label='Validation mae')
plt.title('Training and validation MAE')
plt.legend()

plt.figure()
plt.plot(epochs, loss, 'b', label='Training loss')
plt.plot(epochs, val_loss, 'r', label='Validation loss')
plt.title('Training and validation loss MSE')
plt.legend()

#If we want to reload a trainend model
#NeuNet = tf.keras.models.load_model(save_path)

#INFERENCE ON TEST SET

#Export metrics on CSV

with open(save_path / 'Metrics.csv', "w") as f:
  fieldnames = ['Variable','CC','RMSE','SI','BIAS','MeanAPD','StdAPD']
  writer = csv.DictWriter(f, fieldnames=fieldnames)
  writer.writeheader()


# make predictions on the testing data
preds = NeuNet.predict(x_test_scaled)

for i in range(0,len(targets)):

    name = targets[i]

    pred = preds[i].flatten()
    y_test_var = y_test.iloc[:,i]
```

```
diff = pred - y_test_var
percentDiff = (diff / y_test_var) * 100

#METRICS
AbsPercentDiff = np.abs(percentDiff)
MeanAPD = np.mean(AbsPercentDiff)
StdAPD = np.std(AbsPercentDiff)

CCTest  = pearsonr(y_test_var,pred)   #it should be the same thing of MATLAB 'corr'
RmseTest  = sqrt(mean_squared_error(y_test_var,pred))
SITest  = sqrt(mean_squared_error(y_test_var,pred))/np.mean(y_test_var)
BiasTest  = np.sum(pred-y_test_var)/y_test_var.shape[0]

s = f"""
{'-'*40}
# Variable: {name}
# CC: {CCTest}
# RMSE: {RmseTest}
# SI: {SITest}
# BIAS: {BiasTest}
# MeanAPD: {MeanAPD}
# StdAPD: {StdAPD}

{'-'*40}
"""

print(s)

#Export metrics on CSV

with open(save_path / 'Metrics.csv', "a") as f:
  writer = csv.DictWriter(f, fieldnames=fieldnames)
  writer.writerow({'Variable':name,'CC':CCTest,'RMSE':RmseTest,'SI':SITest,
                  'BIAS':BiasTest,'MeanAPD':MeanAPD,'StdAPD':StdAPD})
plt.figure()
a = plt.axes(aspect='equal')
plt.scatter(y_test_var, pred)
plt.xlabel('Measured '+name)
plt.ylabel('Predicted '+name)
lims = [0, 50]
```


digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

```
    plt.xlim(lims)
    plt.ylim(lims)
    _ = plt.plot(lims, lims)


    plt.figure()
    plt.hist(diff, bins=30)
    plt.title('Error = Target - Predicted for '+name)
    plt.xlabel('Prediction Errors')
    _ = plt.ylabel('Count')


    plt.figure()
    plt.plot(y_test_var, 'ro',markersize=2, markevery=1, marker='h', markeredgecolor='black', markeredgewidth=0.0, label='Meaured '+ name)
    plt.plot(pd.Series(pred,index=y_test_var.index), 'b*',markersize=2, markevery=1, marker='*', markeredgecolor='black', markeredgewidth=0.0,linewidth=1, label='ANN predicted '+name)


    plt.title('Measured V.S. ANN predicted '+name+' on test set' )
    plt.legend()
    plt.xlabel("Data Index")
    plt.ylabel(name)
    #plt.savefig(save_path+ 'Time series full Data.tif', format='tif', dpi=300)
```

Validate ANN Multitask model on laboratory data

```
import os
import sys
import random
import numpy as np
import pandas as pd
from pathlib import Path
import tensorflow as tf
from tensorflow.keras.models import Model
from tensorflow.keras.layers import Input, Dense
from tensorflow.keras.optimizers import Adam, SGD
from tensorflow.keras.callbacks import EarlyStopping
from tensorflow.keras.backend import clear_session
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
import matplotlib.pyplot as plt
from scipy.stats import pearsonr
from sklearn.metrics import mean_squared_error
```


digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

```
from math import sqrt
import csv
from pickle import load

random.seed(2)
random_state = 2


#WORKING DIRECTORY
path = Path('path/to/insert')
save_path = Path(path / 'Ann_model_MLT')


#DATA LOADING
data = pd.read_excel(path / 'Sensor_Lab_datasets_UPDATE_v2.xlsx', header=0)
#data_day = data.groupby(by=data['Date'].dt.date).mean()


feat = ['Influent Flow',
'Influent pH',
'Biofor DN Temperature',
'Biofor CN - Dissolved oxygen',
'Effluent Flow',
'Effluent pH']

targets = list(data.columns[-3:])


cols_to_update = data.columns[:1].tolist() + feat + data.columns[-3:].tolist()
data.columns = cols_to_update

dt = data['Date']
day = pd.Timedelta(1, "h")
#in_block = ((dt - dt.shift(-1)).abs() >= day) | (dt.diff() == day)
breaks_day = dt.diff() != day
breaks_feat = data[['Influent Flow','Influent pH','Effluent Flow','Effluent pH']].diff() != 0
breaks = np.logical_and(breaks_day, breaks_feat.any(axis=1))
data['grp_date'] = breaks.cumsum()
data_day = data.groupby('grp_date').mean()
```


digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

```
x_test = data[feat]
y_test = data[targets]

#MODEL and SCALER LOADING
NeuNet = tf.keras.models.load_model(save_path)
ss = load(open(save_path / 'scaler_MLT.pkl', 'rb'))



#SCALE DATA
x_test_scaled  = ss.transform(x_test)
print(x_test_scaled.shape)
#Export metrics on CSV

with open(save_path / 'Val_Metrics.csv', "w") as f:
  fieldnames = ['Variable','CC','RMSE','SI','BIAS','MeanAPD','StdAPD']
  writer = csv.DictWriter(f, fieldnames=fieldnames)
  writer.writeheader()

#INFERENCE ON TEST SET

# make predictions on the testing data
preds = NeuNet.predict(x_test_scaled)
preds = pd.DataFrame(np.concatenate(preds,axis=1))
preds['Date'] = data['Date']
preds['grp_date'] = breaks.cumsum()
preds_day = preds.groupby('grp_date').mean()

for i in range(0,len(targets)):

    name = targets[i]

    pred = preds[i]
    y_test_var = y_test.iloc[:,i]

    diff = pred - y_test_var
    percentDiff = (diff / y_test_var) * 100

    #METRICS
```


digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

```
AbsPercentDiff = np.abs(percentDiff)
MeanAPD = np.mean(AbsPercentDiff)
StdAPD = np.std(AbsPercentDiff)

CCTest  = pearsonr(y_test_var,pred)   #it should be the same thing of MATLAB 'corr'
RmseTest  = sqrt(mean_squared_error(y_test_var,pred))
SITest  = sqrt(mean_squared_error(y_test_var,pred))/np.mean(y_test_var)
BiasTest  = np.sum(pred-y_test_var)/y_test_var.shape[0]

s = f"""
{'-'*40}
# Variable: {name}
# CC: {CCTest}
# RMSE: {RmseTest}
# SI: {SITest}
# BIAS: {BiasTest}
# MeanAPD: {MeanAPD}
# StdAPD: {StdAPD}

{'-'*40}
"""

print(s)

#Export metrics on CSV

with open(save_path / 'Val_Metrics.csv', "a") as f:
  writer = csv.DictWriter(f, fieldnames=fieldnames)
  writer.writerow({'Variable':name,'CC':CCTest,'RMSE':RmseTest,'SI':SITest,
                'BIAS':BiasTest,'MeanAPD':MeanAPD,'StdAPD':StdAPD})
plt.figure()
a = plt.axes(aspect='equal')
plt.scatter(y_test_var, pred)
plt.xlabel('Measured '+name)
plt.ylabel('Predicted '+name)
lims = [0, 50]
plt.xlim(lims)
plt.ylim(lims)
_ = plt.plot(lims, lims)
```


digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

```
    plt.figure()
    plt.hist(diff, bins=30)
    plt.title('Error = Target - Predicted for '+name)
    plt.xlabel('Prediction Errors')
    _ = plt.ylabel('Count')
```

TSS forecasting

```
import os
import sys
import random
import numpy as np
import pandas as pd
from pathlib import Path
import tensorflow as tf
from tensorflow.keras.models import Model
from tensorflow.keras.layers import Input, Dense
from tensorflow.keras.optimizers import Adam, SGD
from tensorflow.keras.callbacks import EarlyStopping
from tensorflow.keras.backend import clear_session
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
import matplotlib.pyplot as plt
from scipy.stats.stats import pearsonr
from sklearn.metrics import mean_squared_error
from math import sqrt
import csv
from pickle import dump
from sklearn.model_selection import GroupShuffleSplit

clear_session()

def compute_metrics(pred,y_test):

    diff = pred - y_test
    percentDiff = (diff / y_test) * 100

    #METRICS
    AbsPercentDiff = np.abs(percentDiff)
    MeanAPD = np.mean(AbsPercentDiff)
    StdAPD = np.std(AbsPercentDiff)
```


digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

```
    CCTest  = pearsonr(y_test,pred)   #it should be the same thing of MATLAB 'corr'
    RmseTest  = sqrt(mean_squared_error(y_test,pred))
    SITest  = sqrt(mean_squared_error(y_test,pred))/np.mean(y_test)
    BiasTest  = np.sum(pred-y_test)/y_test.shape[0]


    return [CCTest[0],RmseTest,SITest,BiasTest,MeanAPD,StdAPD]


'''fixing seed'''
seed_value = 2

os.environ['PYTHONHASHSEED']=str(seed_value)
random.seed(seed_value)
random_state = seed_value
np.random.seed(seed_value)
tf.random.set_seed(seed_value)

#WORKING DIRECTORY
path = Path('path/to/insert')
save_path = Path(path / 'forecast')



#DATA LOADING
data = pd.read_excel(path / 'data/Biowin simulated data.xlsx', header=1)

#variables to consider
feat = ['Influent Flow',
'Influent pH',
'Biofor DN Temperature',
'Biofor CN - Dissolved oxygen',
'Effluent Flow',
'Effluent pH']

data = data.rename(columns={"Effluent Total suspended solids": "TSS"})
targets = list(data.columns[-1:])

data['new_date'] = [d.date() for d in data['Date']]
dt = data['new_date']
```


digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

```
day = pd.Timedelta(1, "d")
breaks_day = dt.diff() >= day
data['grp_date'] = breaks_day.cumsum()



n = len(data)
train = data[0:int(n*0.7)]
val = data[int(n*0.7):int(n*0.8)]
test = data[int(n*0.8):]

fig, axs = plt.subplots(figsize=(12, 4))
train.plot( 'Date' , targets, ax=axs)
val.plot( 'Date' , targets, ax=axs)
test.plot( 'Date' , targets, ax=axs)
axs.set_title('TSS 1 year timeseries simulation')
axs.legend(['TSS_train', 'TSS_val', 'TSS_test'])
axs.set_xlabel("Date")
axs.set_ylabel('TSS value')
fig.savefig(str(save_path)+ '/TSS timeseries', format='tif', dpi=300,bbox_inches='tight')

fig, axs = plt.subplots(figsize=(12, 4))
test.plot( 'Date' , targets, ax=axs, color='g')
axs.set_title('TSS timeseries simulation test set')
axs.legend(['TSS_test'])
axs.set_xlabel("Date")
axs.set_ylabel('TSS value')
fig.savefig(str(save_path)+ '/TSS test set', format='tif', dpi=300,bbox_inches='tight')

sel = feat+targets
train_df = train[sel]
val_df = val[sel]
test_df = test[sel]

train_mean = train_df.mean()
train_std = train_df.std()

train_df = (train_df - train_mean) / train_std
val_df = (val_df - train_mean) / train_std
test_df = (test_df - train_mean) / train_std
```


digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

```
num_features = train_df.shape[1]


class WindowGenerator():
  def __init__(self, input_width, label_width, shift,
               train_df=train_df, val_df=val_df, test_df=test_df,
               label_columns=None):
    # Store the raw data.
    self.train_df = train_df
    self.val_df = val_df
    self.test_df = test_df

    # Work out the label column indices.
    self.label_columns = label_columns
    if label_columns is not None:
      self.label_columns_indices = {name: i for i, name in
                                    enumerate(label_columns)}
    self.column_indices = {name: i for i, name in
                           enumerate(train_df.columns)}

    # Work out the window parameters.
    self.input_width = input_width
    self.label_width = label_width
    self.shift = shift

    self.total_window_size = input_width + shift

    self.input_slice = slice(0, input_width)
    self.input_indices = np.arange(self.total_window_size)[self.input_slice]

    self.label_start = self.total_window_size - self.label_width
    self.labels_slice = slice(self.label_start, None)
    self.label_indices = np.arange(self.total_window_size)[self.labels_slice]

  def __repr__(self):
    return '\n'.join([
        f'Total window size: {self.total_window_size}',
        f'Input indices: {self.input_indices}',
        f'Label indices: {self.label_indices}',
```


digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

```
      f'Label column name(s): {self.label_columns}'])

OUT_STEPS = 24
wide_window = WindowGenerator(
    input_width=24, label_width=OUT_STEPS, shift=6,
    label_columns=[targets[0]])

def split_window(self, features):
  inputs = features[:, self.input_slice, :]
  labels = features[:, self.labels_slice, :]
  if self.label_columns is not None:
    labels = tf.stack(
        [labels[:, :, self.column_indices[name]] for name in self.label_columns],
        axis=-1)

  # Slicing doesn't preserve static shape information, so set the shapes
  # manually. This way the `tf.data.Datasets` are easier to inspect.
  inputs.set_shape([None, self.input_width, None])
  labels.set_shape([None, self.label_width, None])

  return inputs, labels

WindowGenerator.split_window = split_window

# Stack three slices, the length of the total window.
example_window = tf.stack([np.array(test_df[:wide_window.total_window_size]),
                           np.array(test_df[820:820+wide_window.total_window_size]),
                           np.array(test_df[1200:1200+wide_window.total_window_size])])

example_inputs, example_labels = wide_window.split_window(example_window)

print('All shapes are: (batch, time, features)')
print(f'Window shape: {example_window.shape}')
print(f'Inputs shape: {example_inputs.shape}')
print(f'Labels shape: {example_labels.shape}')

wide_window.example = example_inputs, example_labels

def plot(self, model=None, plot_col=targets[0], max_subplots=3):
  inputs, labels = self.example
```


digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

```
  plt.figure(figsize=(12, 8))
  plot_col_index = self.column_indices[plot_col]
  max_n = min(max_subplots, len(inputs))
  for n in range(max_n):
    plt.subplot(max_n, 1, n+1)
    plt.ylabel(f'{plot_col} [normed]')
    plt.plot(self.input_indices, inputs[n, :, plot_col_index],
             label='Inputs', marker='.', zorder=-10)

    if self.label_columns:
      label_col_index = self.label_columns_indices.get(plot_col, None)
    else:
      label_col_index = plot_col_index

    if label_col_index is None:
      continue

    plt.scatter(self.label_indices, labels[n, :, label_col_index],
                edgecolors='k', label='Labels', c='#2ca02c', s=64)
    if model is not None:
      predictions = model(inputs)
      plt.scatter(self.label_indices, predictions[n, :, label_col_index],
                  marker='X', edgecolors='k', label='Predictions',
                  c='#ff7f0e', s=64)

    if n == 0:
      plt.legend()

  plt.xlabel('Time [h]')

WindowGenerator.plot = plot
wide_window.plot()


def make_dataset(self, data):
  data = np.array(data, dtype=np.float32)
  ds = tf.keras.utils.timeseries_dataset_from_array(
      data=data,
      targets=None,
      sequence_length=self.total_window_size,
```


digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

```
      sequence_stride=1,
      shuffle=True,
      batch_size=32,)

  ds = ds.map(self.split_window)

  return ds

WindowGenerator.make_dataset = make_dataset

@property
def train(self):
  return self.make_dataset(self.train_df)

@property
def val(self):
  return self.make_dataset(self.val_df)

@property
def test(self):
  return self.make_dataset(self.test_df)

@property
def example(self):
  """Get and cache an example batch of `inputs, labels` for plotting."""
  result = getattr(self, '_example', None)
  if result is None:
    # No example batch was found, so get one from the `.train` dataset
    result = next(iter(self.train))
    # And cache it for next time
    self._example = result
  return result

WindowGenerator.train = train
WindowGenerator.val = val
WindowGenerator.test = test
WindowGenerator.example = example_window

wide_window.train.element_spec
```


digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

```
for example_inputs, example_labels in wide_window.train.take(1):
  print(f'Inputs shape (batch, time, features): {example_inputs.shape}')
  print(f'Labels shape (batch, time, features): {example_labels.shape}')



MAX_EPOCHS = 20

def compile_and_fit(model, window, patience=5):
  early_stopping = tf.keras.callbacks.EarlyStopping(monitor='val_loss',
                                                    patience=patience,
                                                    mode='min')

  model.compile(loss=tf.losses.MeanSquaredError(),
                optimizer=tf.optimizers.Adam(),
                metrics=[tf.metrics.MeanAbsoluteError()])

  history = model.fit(window.train, epochs=MAX_EPOCHS,
                      validation_data=window.val,
                      callbacks=[early_stopping])
  return history



multi_lstm_model = tf.keras.Sequential([
    # Shape [batch, time, features] => [batch, lstm_units].
    # Adding more `lstm_units` just overfits more quickly.
    tf.keras.layers.LSTM(32, return_sequences=False),
    # Shape => [batch, out_steps*features].
    tf.keras.layers.Dense(OUT_STEPS*num_features,
                          kernel_initializer=tf.initializers.zeros()),
    # Shape => [batch, out_steps, features].
    tf.keras.layers.Reshape([OUT_STEPS, num_features])
])

history = compile_and_fit(multi_lstm_model, wide_window)
mae = history.history['mean_absolute_error']
val_mae = history.history['val_mean_absolute_error']
loss = history.history['loss']
val_loss = history.history['val_loss']
```


digital-water.city
digitalwater_eu
digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

```
lista = [loss,val_loss,mae,val_mae]
epochs = range(len(loss))

plt.figure()
plt.plot(epochs, loss, 'b', label='Training loss')
plt.plot(epochs, val_loss, 'r', label='Validation loss')
plt.title('Training and validation loss MSE')
plt.legend()



val_performance = multi_lstm_model.evaluate(wide_window.val)
performance = multi_lstm_model.evaluate(wide_window.test, verbose=0)
#wide_window.plot(multi_lstm_model)

predictions = multi_lstm_model.predict(wide_window.test, verbose=0)



res = []
res_round = []

for inputs, labels in wide_window.test:
    labels = np.array(labels[:, :, -1])
    labels = labels*train_std['TSS']+train_mean['TSS']



    pred = multi_lstm_model(inputs)
    pred = np.array(pred[:, :, -1])
    pred = pred*train_std['TSS']+train_mean['TSS']

    for i in range(pred.shape[0]):
        res.append(compute_metrics(pred[i,:],labels[i,:].reshape(-1)))
        res_round.append(compute_metrics(pred[i,:].round(),labels[i,:].reshape(-1).round()))

mean_res = np.mean(res, axis=0)
mean_res_round = np.mean(res_round, axis=0)
```


digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.

Digital Water .City

Leading urban water management to its digital future

digital-water.city
digitalwater_eu

digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.