A set of climate change scenarios and related impacts summarized through infographics, and guidelines for tipping point identification

Deliverable D4.1, WP4

Project Number	**101081980**
Project Title	Climate-resilient management for safe disinfected and non-disinfected water supply systems
Project Acronym	SafeCREW
Project Duration	November 2022 – May 2026
Call identifier	HORIZON-CL6-2022-ZEROPOLLUTION-01
Due date of Deliverable	Month 20, 30.06.2024
Final version date	Month 20, 30.06.2024
Dissemination Level	PU (Public)
Deliverable No.	D4.1
Work Package	WP4
Task	T4.1
Lead Beneficiary	EUT
Contributing Beneficiaries	POLIMI, DVGW-TUHH, KWB, MM, UBA
Report Author	Queralt Plana (EUT), Laura Vinardell (EUT), Beatrice Cantoni (POLIMI), Massimiliano Arca (POLIMI)
Reviewed by	Aleix Martorell (CAT)
Approved by	Mathias Ernst (TUHH)

This project has received funding from the European Union’s Horizon Europe research and innovation programme under grant agreement No 101081980.

	History of Changes		
Version	**Authors**	**Publication Date**	**Change**
0.1	Queralt Plana, Laura Vinardell, Beatrice Cantoni, Massimiliano Arca	31 May 2024	Deliverable preparation for review
0.2	Mathias Ernst, Aleix Martorell	12 June 2024	Revision
0.3	Queralt Plana, Laura Vinardell	14 June 2024	Introduction of reviewers' comments
0.4	Mathias Ernst	21 June 2024	Final review
1.0		27 June 2024	Submitted to Commission

Part. No.	**Participant organisation name**	**Short name**
1 (Coordinator)	German Association for Gas and Water, local research branch at Hamburg University of Technology	DVGW-TUHH
2	Politecnico di Milano	POLIMI
3	Berlin Centre of Competence for Water	KWB
4	BioDetection Systems B.V.	BDS
5	Centre Tecnológic de Catalunya (EURECAT)	EUT
6	German Environment Agency	UBA
7	Helmholtz-Centre for Environmental Research	UFZ
8	Consorci Concessionari d'Aigües per als Ajuntaments i Industries de Tarragona	CAT
9	Tutech Innovation GmbH	TUTECH
10	Metropolitana Milanese SPA	MM
11	Multisensor systems Ltd.	MSS

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

Abstract

Climate change is a setback for water quality. For instance, rising temperatures create ideal conditions for harmful algal blooms, while extreme weather events like floods and droughts disrupt water treatment processes. Floods wash pollutants and sediment into rivers and lakes, making them difficult to purify for drinking. Droughts, on the other hand, concentrate contaminants in shrinking bodies of water.

These challenges highlight the need to adapt our water infrastructure. Upgrading treatment plants to handle fluctuations in temperature and contaminant levels will become crucial. Additionally, there may also be the need to invest in new technologies to remove/mitigate the effects of the harmful algal blooms and other emerging threats. By taking proactive steps, we can ensure clean water supplies for future generations.

Within the SafeCREW project, particularly in task T4.1, a guideline is presented to explain the correlations found between climate change and the impact of it to the water sources to produce drinking water. Firstly, a critical review of existing literature is presented to define the main climatic drivers and their impact on water availability and quality. Complementing the literature review, regional historical data are evaluated to explore any observed trend and correlations between water quantity and quality for Tarragona case study (CS#3) for the available period dataset. Additionally, based on the correlations found and existing projections, an estimation of the water quality variation at different horizons is made for a relevant compound to consider on disinfected drinking water supply system (DWSS) like organic matter.

After the detailed study, it can be concluded that despite climate drivers promote the diminution of the water availability and the deterioration of the water quality according to the literature, more restrictive regulations at local, regional and international levels, or any other anthropogenic activities affecting the watershed (both in terms of water quantity and quality) can have an inverse effect on the water systems.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

Table of contents

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

Abbreviations

GA	Grant Agreement
EC	European Commission
DBPs	Disinfection by-products
DWSS	Drinking water supply system
CS	Case study
UVA254	Absorbance at 254nm

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

1. Introduction

Climate change will affect the availability, quality and quantity of water used for basic human needs challenging the assurance of clean water and sanitation for all as declared at the SDG6 (UNESCO, 2020). As underlined by IPCC reports, there is already an evidence of the diminution of the water availability and a deterioration of the water quality potentially caused by the warming temperatures and the extreme weather events (IPCC, 2014; IPPC, 2022). Consequently, it affects food production and forestry, ecosystem functioning, human and animal health, and compliance with environmental quality standards (concerning any water related directive and regulation) (IPCC, 2014).

In front of these extreme events as a consequence of the climate change, like prolonged droughts, severe storms and flooding, sea level changes, or snowpack depletion, current drinking water supply systems (DWSS) are becoming more vulnerable and can be compromised to meet drinking water regulations (Levine et al., 2016; Li et al., 2014). Thus, conventional systems (i.e. coagulation, filtration disinfection and distribution) or riverbank filtration may be required to be adapted in order to improve their resilience and minimize the vulnerability of the system to the pressure of the climate change (Garnier & Holman, 2019; Ma et al., 2022).

The scientific exploration of the interrelated consequences between climate change, surface and groundwater sources, and drinking water quality is a key point to better understand, adapt and optimize the DWSS (Delpla et al., 2009; Garnier & Holman, 2019; Leveque et al., 2021). However, this is a complex task due to the diversity of relevant factors that can affect the water quality from technical to socio-economic levels such as climate projections, economic growth, land use evolution, policies, etc. (Delpla et al., 2009; Ma et al., 2022).

Furthermore, understanding the limitations of changes in the water cycle are required to improve climate and hydrological models essential to develop and implement adaptation strategies to face climate change (IPCC, 2021). For instance, global and regional climate models include the main climate system components, such as atmosphere, land surface, ocean and sea ice, and their interactions. Then, the impacts of climate change on the water cycle are generally estimated considering a diversity of climate scenarios as output of global climate models and introducing them as input to a hydrological model (IPPC, 2022; Parry, 2007). The weakest point still remains as the linkages and feedbacks between climate and hydrological systems, and the scale mismatch between global climate models and regional or catchment hydrological models (Parry, 2007). Thus, downscaling techniques have to be integrated into hydrological studies (Wilby & Wigley, 1997), despite they remain a source of uncertainty to estimations of water cycle variables such as precipitation, soil moisture, runoff and streamflow at the regional scale from both global climate models or global hydrological models (IPPC, 2022).

In this context, the Earth observation component of the European Union's Space programme, named Copernicus, aims to support a better understanding of our planet and managing the environment sustainably. For it, Copernicus services offer near-real-time data on a global scale based on satellite and in situ observations (like ground, sea, or air stations), and it can also be sued for local and regional needs (Copernicus EU, 2024). While processing and analysing the collected data, Copernicus produces value-added information about observed patterns to elaborate better forecasts using climate and hydrological models.

Based on Copernicus available information, there is a clear proof that climate change will affect surface and groundwater quality and availability (Copernicus EU, 2024). However, the impact on DWSS and water sources still remains unknown and key research areas need to be addressed (Anderson et al., 2023). Additionally, the projections are mainly focused on water quantity (i.e. precipitation, soil moisture, stream flows) rather than water quality (Copernicus EU, 2024). Within the WP4, the SafeCREW project aims to develop guidelines and tools for integrated DWSS management and DWSS upgrades to minimize the overall risk for consumers. For it, to propose optimal interventions for management, redesign, and adaptation of DWSS under climate change scenarios, they must first be defined for different temporal scales. In this context, the goal of Task 4.1 is to define climate change scenarios and their impact on the water quality of drinking water sources. Specifically, this task consists of:

- Review of the state-of-the-art studies on the impact of climate change on the quality of the water sources, specifically on rivers and groundwater as the case studies of SafeCREW project and considered as main water sources in Europe being 88%, more than 60% of rivers, and almost 25% of groundwater (EEA, 2024; WISE, 2024).
- Exploration of historical regional data and correlations found in the historical data between climate change drivers and drinking water quality variables.
- Identifying relevant climate scenarios and projections according to the IPCC Assessment reports, such as AR5 and AR6, and collect Copernicus projection data on climate change parameters.
- Combination of historical correlation models and climate projections to estimate water quality parameters from climate scenarios.

The consolidation of the existing literature, available information, and projections on Copernicus at regional and global levels, and historical data at a specific case, allow to estimate and define climate scenarios under defined conditions. Additionally, these scenarios can support the decision-making for climate adaptation of the future DWSS.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

2. State of the art on climate change impact on water quality

Climate change has become inevitable and countries in Europe must start (if not started yet) to adapt their water management and treatment systems to minimise their vulnerability. A general review of the existing literature to identify the main impact of climate change on drinking water sources from worldwide to regional levels to identify the tipping point for water system adaptation is presented in this chapter.

The interconnections between the climatic drivers and their impact on water systems based on literature review are shown in Figure 1. At worldwide level, the temperature during the last decade has already increased 1.1 ºC and it will get worse in the coming decades (IPCC, 2021). As a consequence of the air temperature raise, the evaporation and evapotranspiration are also favoured (C. D. Evans et al., 2011; M. Evans et al., 2006; IPPC, 2022), thus the soil moisture and groundwater availability will also be affected by this phenomena reducing their water volume (IPCC, 2014; IPPC, 2022), hence the intake to surface water (Mitchell & Likens, 2011). Despite the most noteworthy impacts of climate change are expected to affect surface water quantity and quality, there is a potential decrease of the groundwater quantity and quality expected due to the interdependence with the precipitation and surface water bodies (Amanambu et al., 2020; Kumar, 2012).

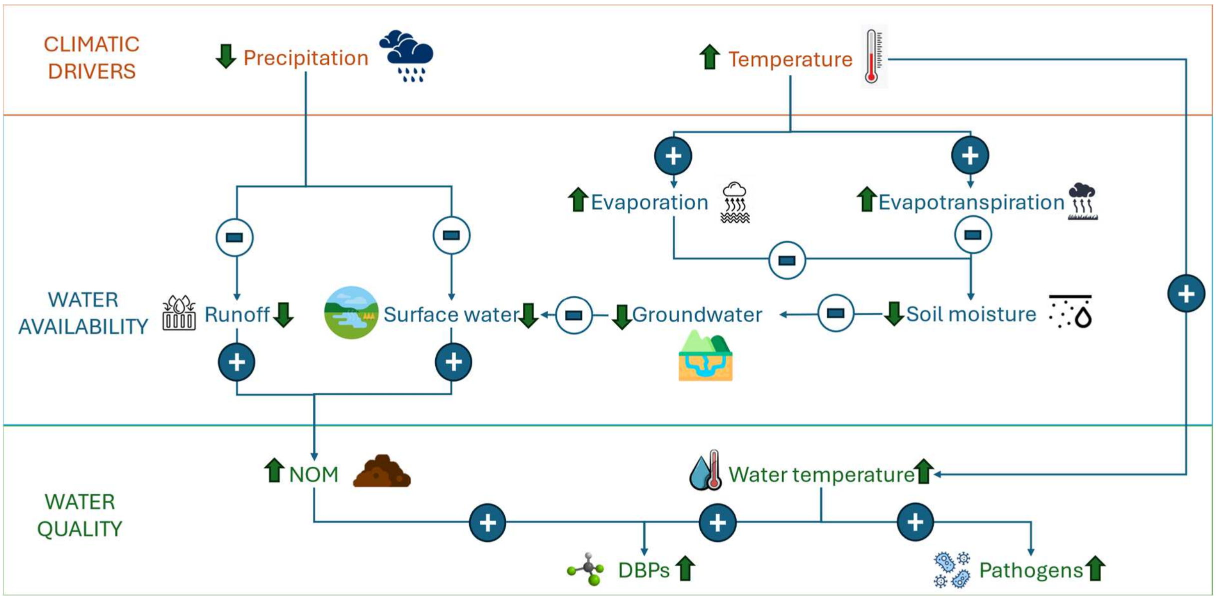

Figure 1. Climate drivers and their impact on water systems.

The precipitation has a direct impact on the surface water quantity and on the runoff occurring during rain events. In Europe, it is expected that the annual precipitation will decrease hence, the contribution to surface water bodies and runoff will also decrease (IPPC, 2022). However, altered patterns increase floods due to the growth of intensity and frequency of severe storms, and extend drought periods (Bangash et al., 2013). Therefore, when a severe storm event occurs, it will increase the runoff and the erosion of the soil as intake of surface water bodies contrary of what is shown in Figure 1 (Delpla et al., 2009).

In general terms, water availability will decrease both as surface water and groundwater, and together with the climatic drivers, it will impact the water quality. For instance, the natural organic matter (NOM) concentrations are predicted to raise, although the composition and the amount of increase are unknown (Creed et al., 2018). Considering the altered precipitation patterns (i.e. severe storm and longer droughts), a greater seasonality in dissolved organic matter (DOM) export is also expected (Ritson et al., 2014), affecting the allochthonous DOM proportion at the water column (Creed et al., 2018). This negative correlation between water availability and NOM can also be observed among

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

other water quality parameters such as dissolved oxygen (DO), conductivity, alkalinity, and other ions (Delpla et al., 2009; Prathumratana et al., 2008).

Furthermore, the increase of the air temperature, has a direct impact on the water temperature which promotes the microbial activity (Couture et al., 2012; Delpla et al., 2009; Finstad et al., 2016; Hofstra, 2011; Lipczynska-Kochany, 2018; Xiao et al., 2020). This is especially relevant for the drinking water production on non-disinfected DWSS. In addition, the disinfection by-products (DBPs) concentrations are also assumed to increase due to the increase of NOM concentrations (DBP precursor) and water temperature that boosts DBPs generation rate (Cool et al., 2019; Delpla & Rodriguez, 2014). This fact is a key factor for the future adaptations for disinfected DWSS.

These changes on the water properties as presented above can lead to a modification of the species diversity. Both, they can promote an increase of the frequency of algal blooms (Ritson et al., 2014). Currently, most of the drinking water treatment plants (DWTP) are not built to remove components as salts or harmful compounds derived from algal blooms. Thus, special attention also needs to be paid to it in order to ensure safety water and minimise adverse health effects (Kouakou & Poder, 2019; Moore et al., 2008; Vineis et al., 2011).

A selection of 22 scientific publications based on European and North American studies was summarized in Figure 2. This graph shows the number of articles, published in the last 20 years, that dealt with the impact of climate change on drinking water quality variables.

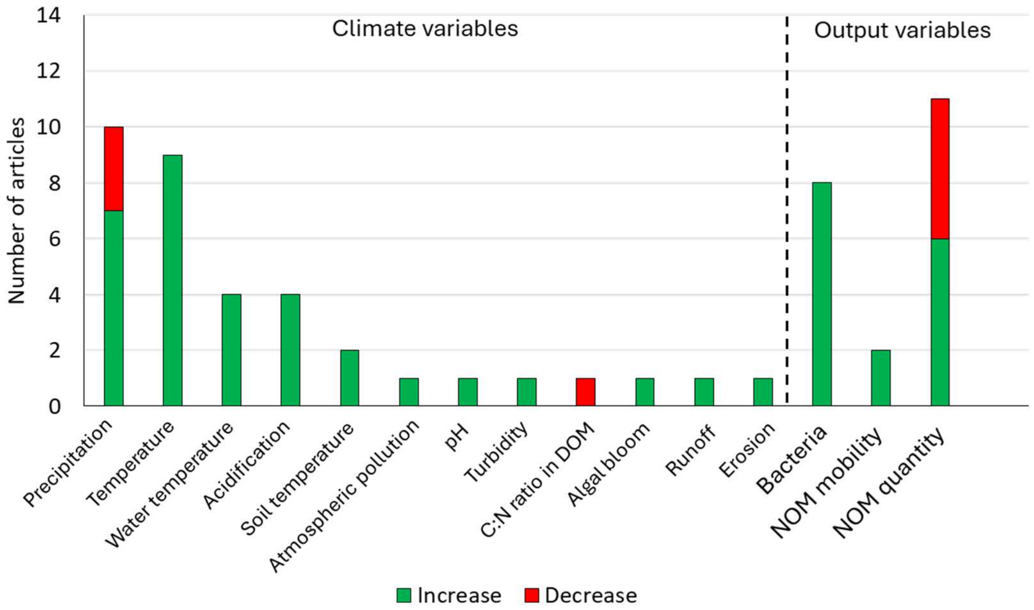

Figure 2. Number of articles reporting the effect of climate change on climate variables and output drinking water quality variables. Bars colours report whether the study found increasing (green) or decreasing (red) trend.

Concerning the climate variables, the most studied were the precipitation and the temperature considered as main climatic drivers, with 10 and 9 studies respectively, while all the rest of the papers on the other have been assessed only in 1 to 4 articles. For all the parameters but the precipitation, only one type of trend is expected per variable: for instance, climate change will probably cause an increase of temperature and acidification (Clark et al., 2005; IPCC, 2014; IPPC, 2022). On the other hand, precipitation will have altered patterns (Bangash et al., 2013; Ritson et al., 2014), showing both an increase in intensity and frequency of storms (Anderson et al., 2023; Clark et al., 2007; Creed et al., 2018; De Wit et al., 2016) and lower yearly rainfall with longer periods of drought (Clark et al., 2005; Strock et al., 2016; Watmough et al., 2016). Consequently, these altered patterns will affect runoff and surface water availability as presented above. For instance, runoff will increase during severe storm episodes as remarked by Delpla et al. (2009).

This project has received funding from the European Union’s Horizon Europe research and innovation programme under grant agreement No 101081980.

Concerning the output drinking water quality variables, bacteria and NOM mobility will raise because of the effects of climate change (Hofstra, 2011; Lipczynska-Kochany, 2018; Xiao et al., 2020), mainly due to the increase of water temperature (C. D. Evans, 2005; Xiao et al., 2020), while there are contrasting predictions concerning NOM quantity: indeed, even if the increase of water temperature is expected to produce a higher release of NOM, it will also promote bacterial activity, causing NOM biodegradation (Delpla et al., 2009; Hofstra, 2011; Sardans et al., 2008). This graph highlighted that it is not clear yet how NOM amount in surface water will change due to climate impacts. Hence, there is a need to obtain monitoring data from long term campaigns (decades), including both quantitative and qualitative information of NOM. Such activities would be of valuable support to predict the effects of climate change on this variable and its implications on drinking water treatment.

Overall, climate change presents a major challenge for water security, especially due to the water availability, increase of bacterial content and changes in NOM concentration. The altered patterns expected ranging between two extreme situations (long drought periods combined with severe storms), DWSS should be built as flexible as possible to be able to treat water in order to ensure safe water according to the respective and current legislative requirements. Researchers are actively studying these impacts and developing adaptation strategies, such as improved water management practices and infrastructure development.

3. Study site and exploration of historical data

3.1 Introduction to the study site

Consorci d’Aigües de Tarragona (CAT) is the institution capturing, treating, and distributing drinking water to 69 municipalities and 25 industries from Tarragona province (CAT, 2024). Raw water comes from the Ebro River channels starting from a diversion dam in Xerta (Figure 3). The water intake is close to the delta (Raw Water Pump Station (EB-0) in Figure 5). The Ebro River begins in the north of Spain (Reinosa, Cantabria), crosses 10 provinces along more than 900km, and ends in Tarragona, Catalonia. Figure 3 shows the Ebro watershed, including the dams (names in black) and Ebro River and its effluents (names in blue). *Confederación Hidrográfica del Ebro* (CHE) is the organism responsible for hydrological planning and water management of this watershed (CHE, 2024).

This project has received funding from the European Union’s Horizon Europe research and innovation programme under grant agreement No 101081980.

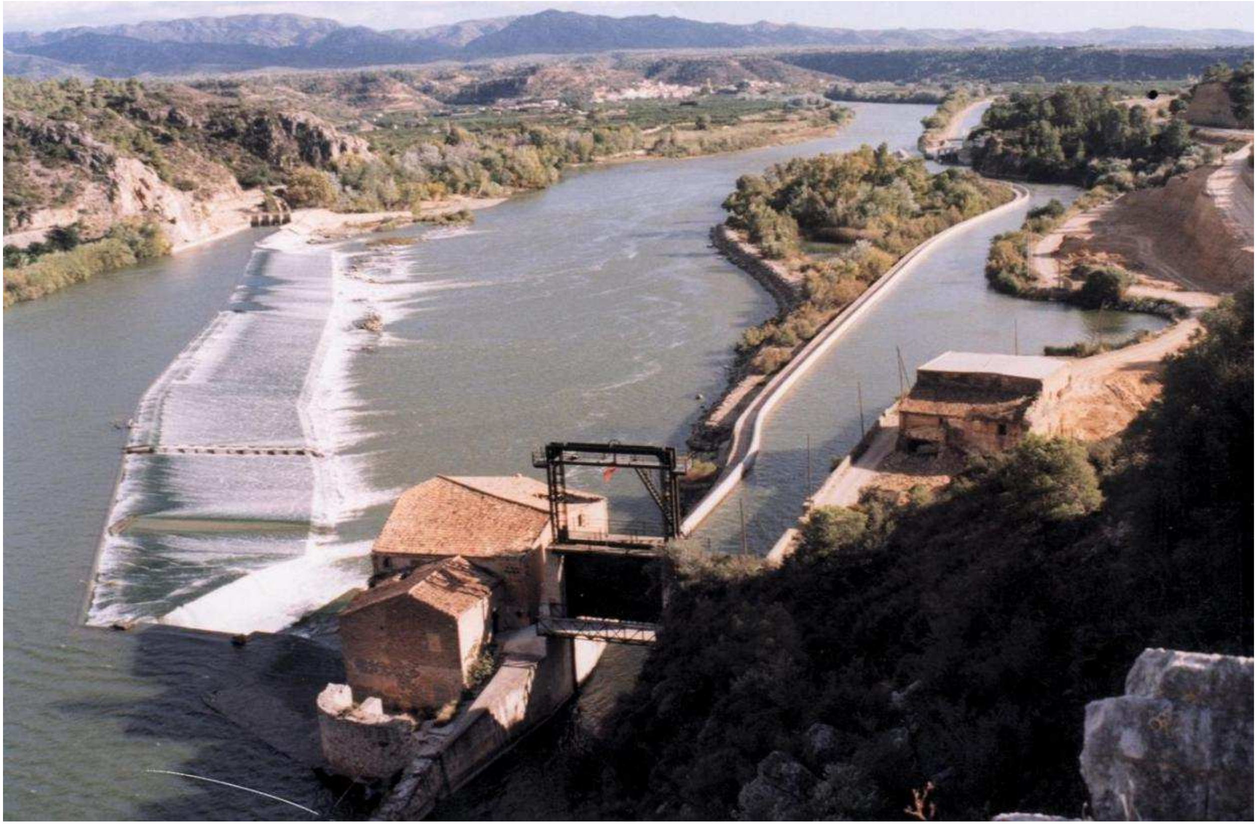

Figure 3: Ebro diversion dam at Xerta: left and right Ebro channels beginning.

Figure 4. Ebro River basin. Source Wikipedia (2024) (under Creative Common License)

3.2 Historical data

The available public datasets can be obtained from *Sistemas Automáticos de Información Hidrológica* (SAIH), and from CAT (CAT, 2024; SAIH Ebro, 2024). SAIH is the tool used by CHE, where quality and meteorological data is stored. The closer CHE stations to CAT water intake (EB0 in Figure 5) are Mequinenza, Guiamets (meteorological data), Xerta and Tortosa (water quality data). Available

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

measures from each station are summarized in Table 1. Data frequency goes from 15min to daily means (user decision).

Figure 5. CHE monitoring stations and CAT water intake (EB0) and DWTP (EB1).

Table 1. Historical data from SAIH and available period.

Measurement	GUIAMETS	MEQUINENZA	TORTOSA	XERTA
Cumulated rainfall	1997 - 2023	1997 - 2023	2004 - 2023	
Environmental temperature	2008 - 2023			
Water temperature			2012 - 2023	1997 - 2023
Conductivity			2012 - 2023	1997 - 2023
River flow			2012 - 2023	
Turbidity			2012 - 2023	1997 - 2023
Absorbance at 254 nm				2012 - 2023
Ammonium				1997 - 2023
Dissolved oxygen				1997 - 2023
Nitrate				2012 - 2023
pH				1997 - 2023
Redox potential			2012 - 2023	

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

3.3 Materials and Methods

The period of analysis was selected based on the target variable, which is, in this study, absorbance at 254nm (1/m) (UVA254) as an indicator of natural organic matter present in the river. Table 1 reports the stations and years for which data were available for each variable. Since UVA254 (1/m) was measured from 09/2012 to 09/2023, this time range was considered for all the other variables.

Since anomalous, duplicate, and missing values in the datasets can significantly affect the validity and reliability of the analyses conducted, a set of pre-processing steps was performed.

a. Determination of anomalous values

For each variable, anomalous values were determined through the following criteria, based on expert knowledge, across the available variables:

- **UVA254 (1/m):** In accordance with a thorough visual examination of the accessible data, a threshold of 60 1/m was chosen, delineating a range beyond which the data was considered to exhibit anomalous characteristics.
- **Air Temperature (°C)**: the range of values was valid.
- **Ammonium (mg/L)**: the EU (Kristensen, 1996) analysed the average ammonium concentration for the biggest rivers in Europe, with an average value for the Ebro River of 0.3 mg/L. So, it was decided to take 1 mg/L as the upper threshold for this measure.
- **Conductivity (µS/cm)**: Looking at Tortosa time series, an unusual value was identified at the beginning of the time series. Therefore, an acceptable threshold range from 250 to 2500 µS/cm was chosen. While, for Xerta, the range of values was proved to be valid.
- **Daily Cumulated Rainfall (mm/day):** Copernicus (Lopez-Bustins et al., 2019) asserts that in many studies a rainfall is defined as torrential if the daily cumulated rain is greater than 100 mm, while is defined as extreme if greater than 200 mm. Being the data values below the 200 mm threshold, no removal was required.
- **Dissolved Oxygen (mg/L)**: compared to other European rivers (Blaszczak, 2023), the range of values was considered valid.
- **River flow Rate (m^3/s)**: the range of values was considered valid.
- **Nitrate (mg/L)**: the nitrate concentration in the Ebro River showed evidence that the range of measured values was valid (EU, Knowledge Hub for Water), with minimum measures around 5 mg/L and maximum of 39 mg/L.
- **pH**: the range of values was considered valid.
- **Redox Potential (mV)**: the range of ORP can vary from -400 to 800 mV usually (Sigg, 2000), so the value range was considered valid.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

- **Turbidity (NTU):** high levels of turbidity are above 100 NTU (DataStream Initiative, 2021). Levels of turbidity above 100 NTU are unsafe for most aquatic life (In-Situ, 2022). The Ebro River has a wide ecosystem (Wikipedia). Given these considerations, it was decided to consider valid the range of values between 0 and 150 NTU.

- **Water Temperature (°C):** Looking at Tortosa time series, an unusual value was identified at the beginning of the time series. Therefore, an acceptable upper threshold of 35°C was chosen. While, for Xerta, the range of values was proved to be consistent.

b. Similarity analysis between different sampling stations

The target variable measurements were available in Xerta station. A similarity analysis between different stations was done to include into the Xerta dataset (i.e., the merged dataset) those variables only present in other stations. The variables available for more than one station (i.e., conductivity, daily cumulated rainfall, turbidity, and water temperature), were analysed to evaluate similarity between the different sampling points. The metrics used to evaluate the similarity between measurements of the same variable for different stations were the Pearson Correlation Coefficient and the Root Mean Squared Deviation (RMSD).

As for the Pearson Correlation Coefficient, let ***X*** and ***Y*** be two datasets of the same variable measured in two stations with *n* samples each and $x_i \in$ ***X***, $y_i \in$ ***Y***, *i* = {1, …, n} be the i^th samples for variables ***X*** and ***Y***, respectively. Then, the Pearson Correlation Coefficient is defined as:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=0}^{n}(y_i - \bar{y})^2}},$$

where $\bar{x}$ and $\bar{y}$ are the expected values (i.e., mean value) of the ***X*** and ***Y*** datasets, $r \in [-1, 1]$. In particular, if such a coefficient approaches either 1 or -1, it indicates that when a variable changes, the other variable changes in the same (or opposite) direction, respectively. Given that the comparison involves measurements of the same variable across different stations, if the stations are homogeneous, the expected coefficient r will show a strong and positive relationship, so r approaching a value of 1.

A Pearson correlation matrix displays the Pearson correlation coefficients for each pair of variables, measuring their linear relationships. Each cell in the matrix contains a coefficient ranging from -1 to 1. A value of 1 indicates a perfect positive linear relationship, -1 indicates a perfect negative linear relationship, and 0 indicates no linear relationship. Diagonal elements are always 1, showing each variable's correlation with itself. The matrix helps identifying the strength and direction of relationships between variables, aiding in exploratory data analysis, feature selection in machine learning, and identifying multicollinearity in regression analysis.

The Root Mean Square Deviation (RMSD) is defined as:

$$RMSD = \sqrt{\frac{\sum_{i=1}^{n}(x_i - y_i)^2}{n}}.$$

It is a measure of distance in the value of the two datasets. Therefore, the lower the metric, the more similar the two variables are.

This process was performed for both the conductivity and water temperature variables, while different considerations were made for turbidity and daily cumulative rainfall. More specifically, turbidity measurements were taken directly from Xerta due to the significant number of missing values in the ones from Tortosa. On the other hand, daily cumulated rainfall measurements were selected from Tortosa, as it is the closest station to Xerta.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

Once the final dataset was built, it was refactored by computing the monthly average for each variable to be compliant with the available projections provided by the Copernicus project.

c. Moving average calculation

To better understand the trend of each variable and the relationships with respect to the UVA254, the trend was analysed and computed through the Moving Average method. More generally, a moving average of order *m* can be written as:

$$\hat{T}_t = \frac{1}{m} \sum_{j=-k}^{k} y_{t+j},$$

where *m = 2k + 1.* According to this formula, the estimate of the moving average at time *t* is obtained by averaging values of the time series in a defined period centred in time *t*. Given that each variable was averaged monthly and exhibits annual seasonality, it was decided to calculate the moving average at time *t* taking monthly averaged data from 6 months before to 6 months after time *t* (*k* = 6).

d. Correlation analysis

Since Copernicus projections were available for air temperature, monthly cumulated rainfall, river flow, and water temperature, to be able to project absorbance 254nm, it was necessary to study a correlation between UVA254 and the variables for which Copernicus projections were available. A year-wise analysis on the correlation between the previously mentioned variables and the absorbance at 254nm was performed on the historical available data, to understand if there was consistency in the relationships between them through the year.

The metric used to assess the correlation is the Spearman's Rank Correlation Coefficient, which is a technique that can be used to summarize the strength and direction (negative or positive) of a non-linear relationship between two variables. The coefficient is defined as follows:

$$\rho = 1 - \frac{6\sum_{i=1}^{n} {d_i}^2}{n(n^2-1)},$$

where d_i is the difference between ranked observations of the two variables. As for the Pearson Correlation Coefficient, $\rho \in [-1, 1]$. The closer to 1, the stronger and more positive direction the correlation. Conversely, the closer to -1, the stronger and more negative direction the correlation.

For completeness, the relationships between the absorbance at 254nm and the variables with available projections were analysed and the results were compared with the literature. Each variable was split into four groups based on its quartiles (<25th percentile; 25th-50th percentile; 50th-75th percentiles; >75th percentile). The corresponding groups of data points within each quartile's range were analysed to understand if there was statistical evidence of a positive or negative relationship between pairs of variables.

3.4 Exploratory Data Analysis

Anomalous values were removed for each variable according to the rules given in Section 3.3. Table 2 provides a recap of the data preprocessing phase, including the missing data analysis, the anomalous data removal, and the range of available data. Although most of the parameters had datasets before 2012, only the data from 2012 were used since the absorbance at 254nm (UV254), i.e., the target variable to be predicted, was available only from that year.

The Xerta site is the one with the largest number of available variables, while the Mequinenza site is the one with the least. Mequinenza was discarded because the sole available data was rainfall, which was acquired from the Tortosa site, the nearest location.

On average, the percentage of missing values per variable is around 10%. The only variable that has almost half of the missing values is the turbidity at the Tortosa station. Moreover, since all the water quality parameters for the Tortosa site have incoherent values, i.e., most of them being peak values way out of range, there is a possibility that some of the Tortosa site sensors were compromised.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

Table 2. Summary of historical data preprocessing.

		GUIAMETS				**MEQUINENZA**				**TORTOSA**				**XERTA**			
		Historical data				**Historical data**				**Historical data**				**Historical data**			
	Variables	**N° Samples**	**% Missing**	**N° Anomalous**	**Years**	**N° Samples**	**% Missing**	**N° Anomalous**	**Years**	**N° Samples**	**% Missing**	**N° Suspicious**	**Years**	**N° Samples**	**% Missing**	**N° Suspicious**	**Years**
Weather	Air temperature	5498	1.40%	0	2008-2023	-	-	-	-	-	-	-		-	-	-	-
	Daily accumulated precipitation	5498	1.40%	0	2008-2023	9467	1.24%	0	1997-2023	7028	0.43%	0	2004-2023	-	-	-	-
Water quantity	Flow	-	-	-	-	-	-	-	-	4135	1.11%	0	2012-2023	-	-	-	-

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

Water quality	Water temperature	-	-	-	-	-	-	-	-	4066	2.66%	13	2012-2023	9701	8.03%	0	1997-2023
	Turbidity	-	-	-	-	-	-	-	-	4082	44.86 %	74	2012-2023	9700	11.25 %	0	1997-2023
	pH	-	-	-	-	-	-	-	-	-	-	-	-	9761	8.59%	0	1997-2023
	Conductivity	-	-	-	-	-	-	-	-	4081	8.11%	14	2012-2023	9701	8.69%	0	1997-2023
	Dissolved oxygen	-	-	-	-	-	-	-	-	-	-	-	-	9701	11.39 %	0	1997-2023
	UV254	-	-	-	-	-	-	-	-	-	-	-	-	4227	13.30 %	2	2012-2023
	Ammoniu m	-	-	-	-	-	-	-	-	-	-	-	-	9761	13.47 %	7	1997-2023

This project has received funding from the European Union’s Horizon Europe research and innovation programme under grant agreement No 101081980.

	Nitrate	-	-	-	-	-	-	-	-	-	-	-	-	4227	4.16%	0	2012-2023
	Redox Potential	-	-	-	-	-	-	-	-	-	-	-	-	4177	6.58%	0	2012-2023

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

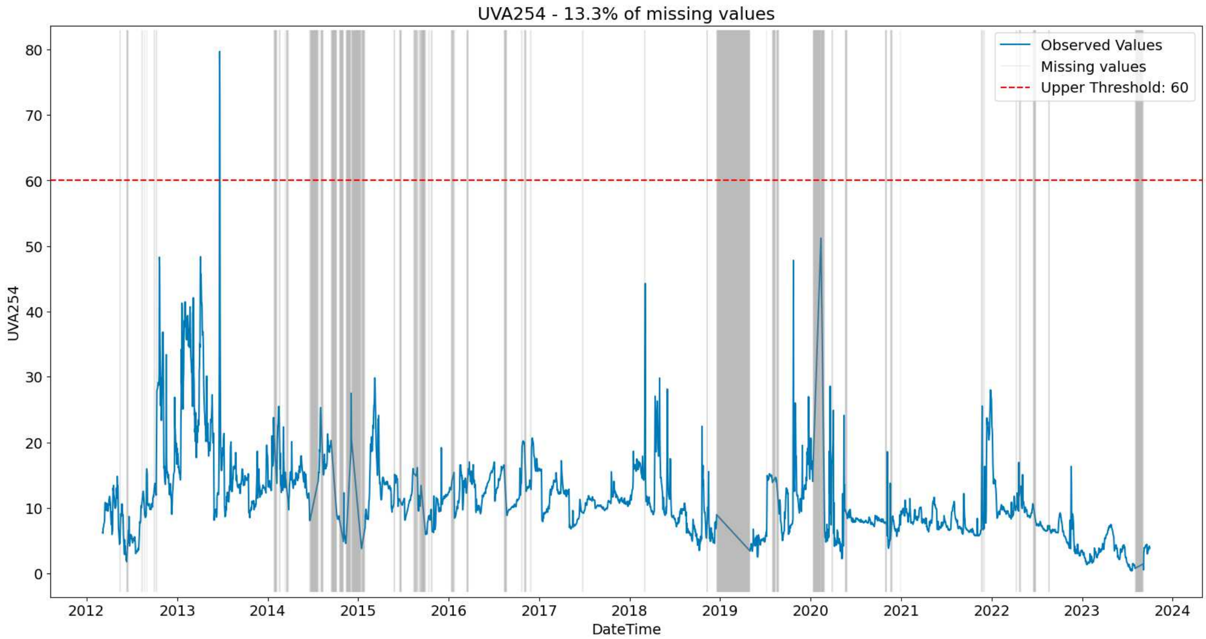

Figure 6. Time series of absorbance at 254 nm observed value. Missing data are highlighted with vertical grey lines. The red dotted line represents the upper threshold imposed to identify and eliminate anomalous data.

Figure 6 shows an example of anomalous data removal for the UV254. In this specific case, the upper threshold was 60 1/m, and only a single data in 2013 was deemed anomalous. The picture also highlights the periods in which data were missing with vertical grey bars. This shows that there were only a few periods longer than 5 months for which there are no missing data. This suggests that the reconstruction phase was necessary for this data set.

To build a unique dataset, it was necessary to assess similarities between the Xerta site and the other ones to incorporate the parameters from the latter to Xerta. Similarities between Xerta and Tortosa were assessed by comparing their common variables (see Figure 7 and Figure 8) and their interactions (see Figure 9). To be able to also retrieve the air temperature available for the Guiamets site, a comparison between the Tortosa and the Guiamets daily cumulative rainfall measurements was performed (see Figures in Annex 1).

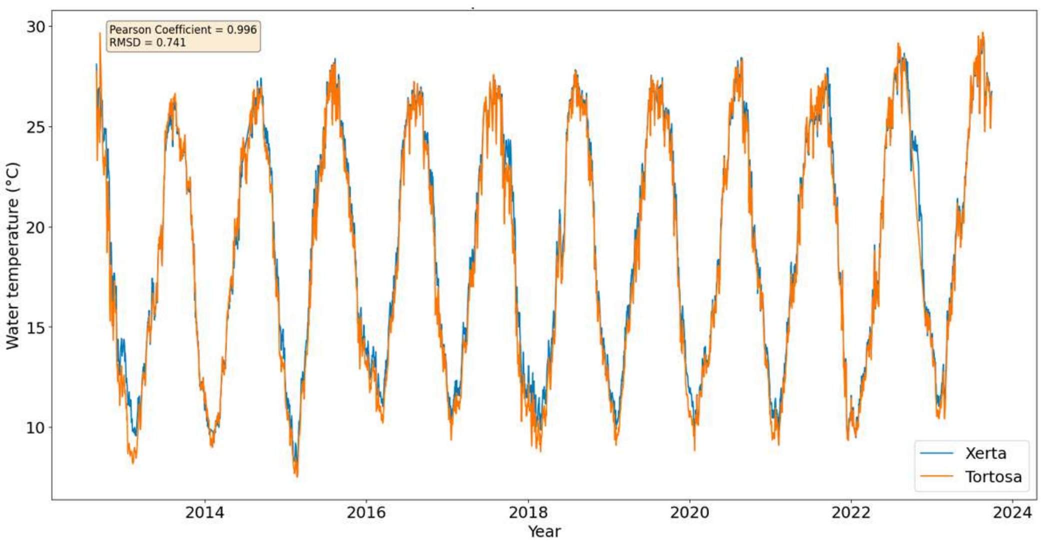

Figure 7. Comparison of time trends for Xerta and Tortosa for water temperature measurements.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

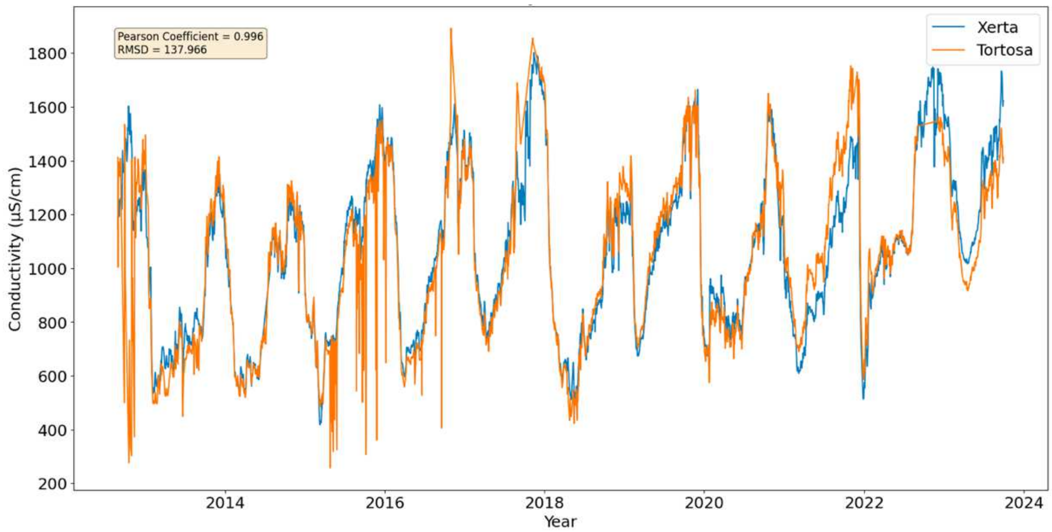

Figure 8. Comparison of time trends for Xerta and Tortosa for conductivity measurements.

In Figure 9, the only two variables presenting significant correlations between Tortosa and Xerta were water temperature and conductivity. In water temperature, there is a high Pearson coefficient (0.996) and low RMSD (0.741), meaning they have similar values and trends. For the conductivity, there is a high Pearson Coefficient (0.908) but a high RMSD (137.966), probably due to the high noise in the measurements for Tortosa, the variable's high order of magnitude and the hydraulic time between both stations.

Regarding the daily cumulated rainfall, the measurements were taken directly from the Tortosa station, which is the closest station to Xerta. However, for completeness, it was performed a comparison between the two variables by taking only those days where both the measurements were greater than 0. The comparison led to a Pearson coefficient of 0.519 and a RMSD of 10.041, that for a parameter with high variability as the rainfall is sufficient to say the variables are similar.

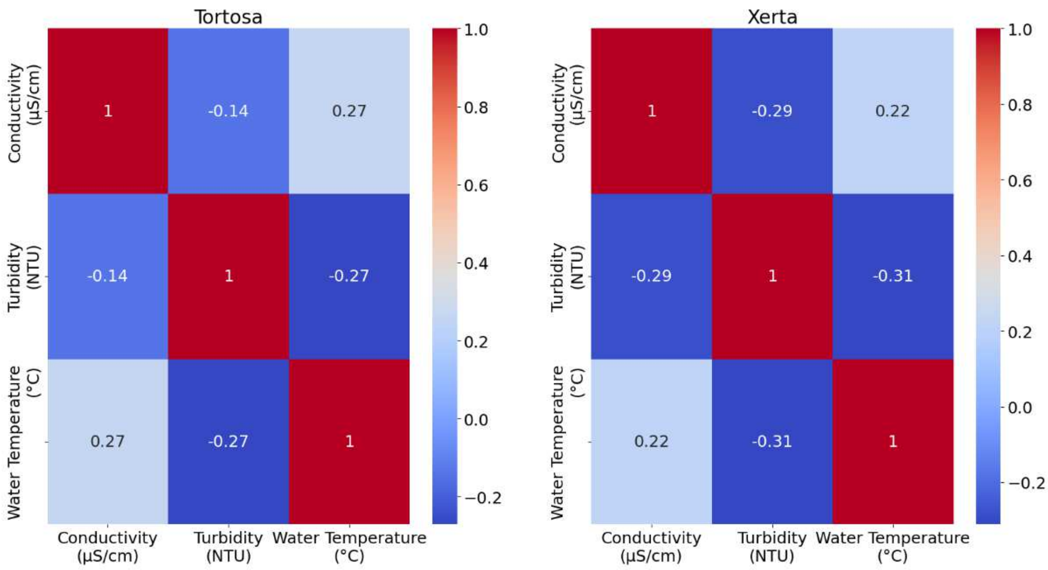

Figure 9. Pearson Correlation matrices between variables from Tortosa and Xerta.

Focusing on the Tortosa and Xerta stations, the Pearson correlation matrix, is shown in Figure 10. The correlations between variables are similar across the two stations. The positive/negative relationship is consistent, while the coefficient values are quite noisy, probably due to the interpolation process for the missing values and the noisy measurements from the Tortosa sensors.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

These results confirm the possibility to incorporate data from different stations onto the Xerta database to create a new unified larger database on which all the further analyses have been performed.

3.5 Time and Trend Analysis

This section reports historical trends for the available variables (air temperature, water temperature, river flow, conductivity, daily cumulated rainfall, turbidity, pH, redox potential, and absorbance at 254nm). The monthly average time series and the corresponding moving average calculated on a 12-month period have been studied for the available variables. Conductivity, air and water temperature variables present an increasement trend. In contrast, river flow, pH, turbidity, redox potential, dissolved oxygen, ammonia (see annex) and UV at 254 nm are decreasing from 2012.

Temperature rise is consistent with literature on climate change impact (see Figure 10): greenhouse gases emission increment implies a generalized higher value of environmental temperature and in consequence of water temperature. River flow decrease is predictable due to the highlighted water scarcity in the Mediterranean region (as Tarragona) reported in IPCC AR5 and AR6, consequence of climate change and land use in arid and semi-arid regions (see Figure 11). Conductivity increase is probably a side effect of the river flow decrease: salts are more concentrated. In a future, it is expected to become worse (see Figure 12). Actions to protect the environment and to ensure safe drinking water for human needs should be considered and planned.

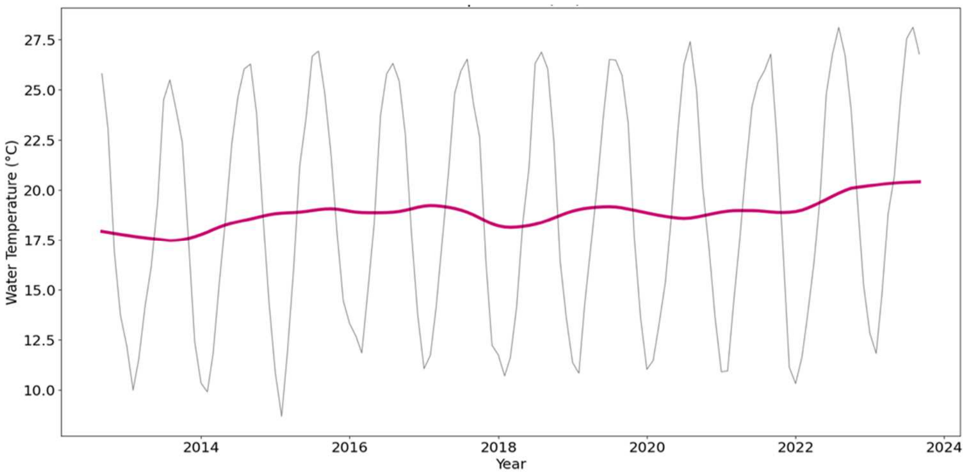

Figure 10. Monthly average time series (black line) and the corresponding moving average (red line) for water temperature.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

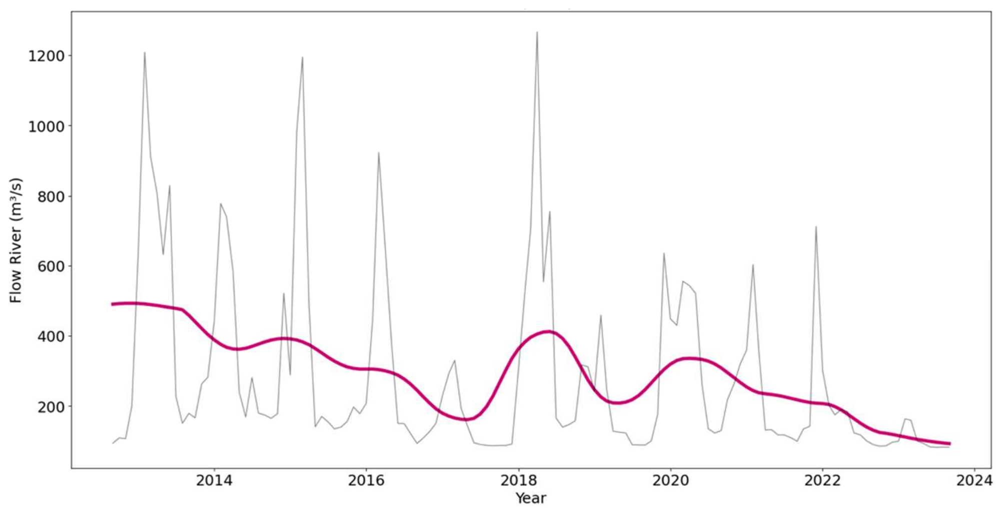

Figure 11. Monthly average time series (black line) and the corresponding moving average (red line) for river flow.

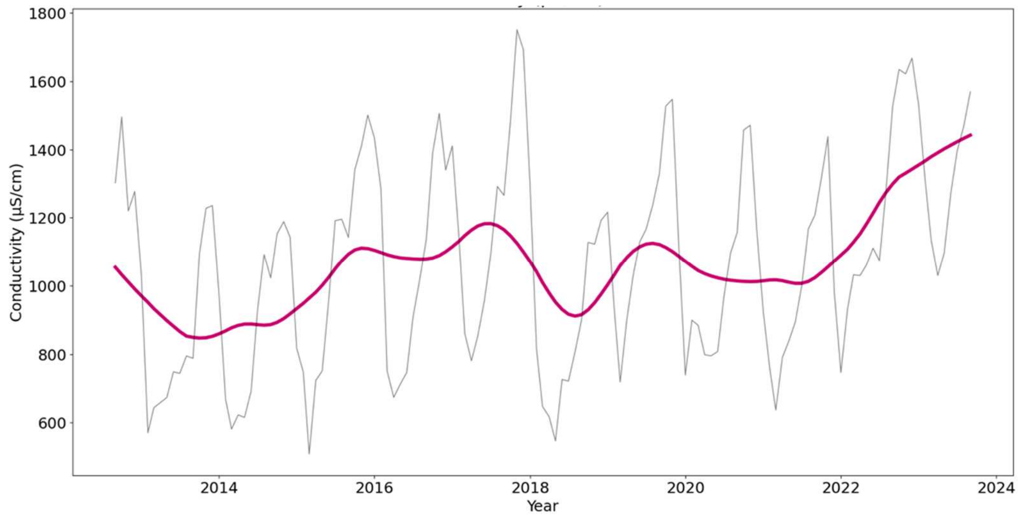

Figure 12. Monthly average time series (black line) and the corresponding moving average (red line) for conductivity.

What is inconsistent with respect to this worsening scenario that is expected with climate change is the absorbance at 254nm decreasing trend. This result highlights a general decreasing trend in organic matter concentration in the monitored river. Typically, human activities like agriculture, industry, and wastewater discharge introduce significant amounts of ammonium and organic matter into water bodies. Therefore, despite climate drivers promoting the deterioration of the water quality with an increase in NOM content according to the literature, the lack of rainfall during the last 3 years led to a runoff decrease, affecting turbidity and NOM measures. In addition, more restrictive regulations on anthropogenic activities at local, regional, and international levels, or any other actions activities such as watershed management may have the opposite effect on the water quality bodies. This could be confirmed by a further study of the decrease in ammonium concentration, turbidity, oxygen demand that follows similar trend than absorbance.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

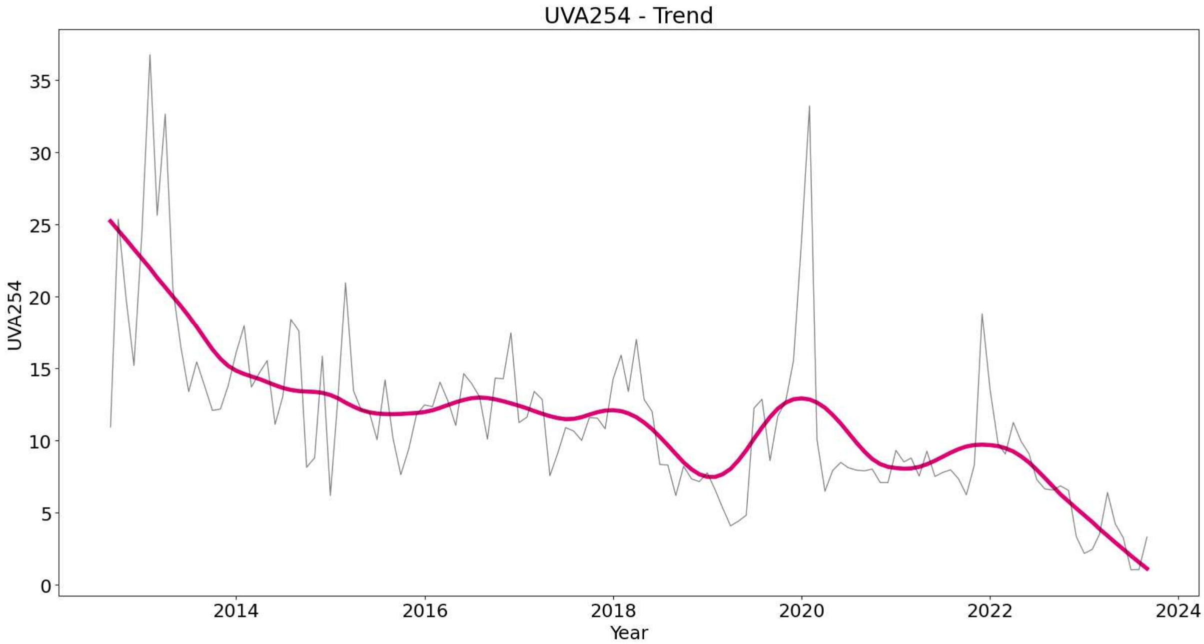

Figure 13. Monthly averaged time series (black line) and the corresponding moving average (red line) for absorbance at 254nm.

The rest of the figures for the variables mentioned above, are presented in annex 1.

Figure 14 presents the Spearman correlation coefficients (described in section 3.3d) of the absorbance at 254nm with air temperature, daily cumulated rainfall, river flow, and water temperature year by year. The only variable showing a consistent negative relationship is water temperature, while the others have different behaviours over the years. The daily cumulated rainfall presents positive correlations for most of the years (stronger in 2012, 2016, 2017, 2022 and 2023). As for the air temperature parameter, on average it is visible that there are no strong correlations except for the years 2017 and 2023. The other years present a weak correlation, symptom of the fact the variable is not quite influent on the variability of the absorbance. A positive significant relationship can be inferred in 2014, 2015, 2019, 2021, 2022 and 2023 for the river flow variable. On the other hand, negative relationships are presents for the 2016 and 2018 years. Overall, it seems hard to find a consistent correlation between the 4 measured variables and the absorbance except for the water temperature.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

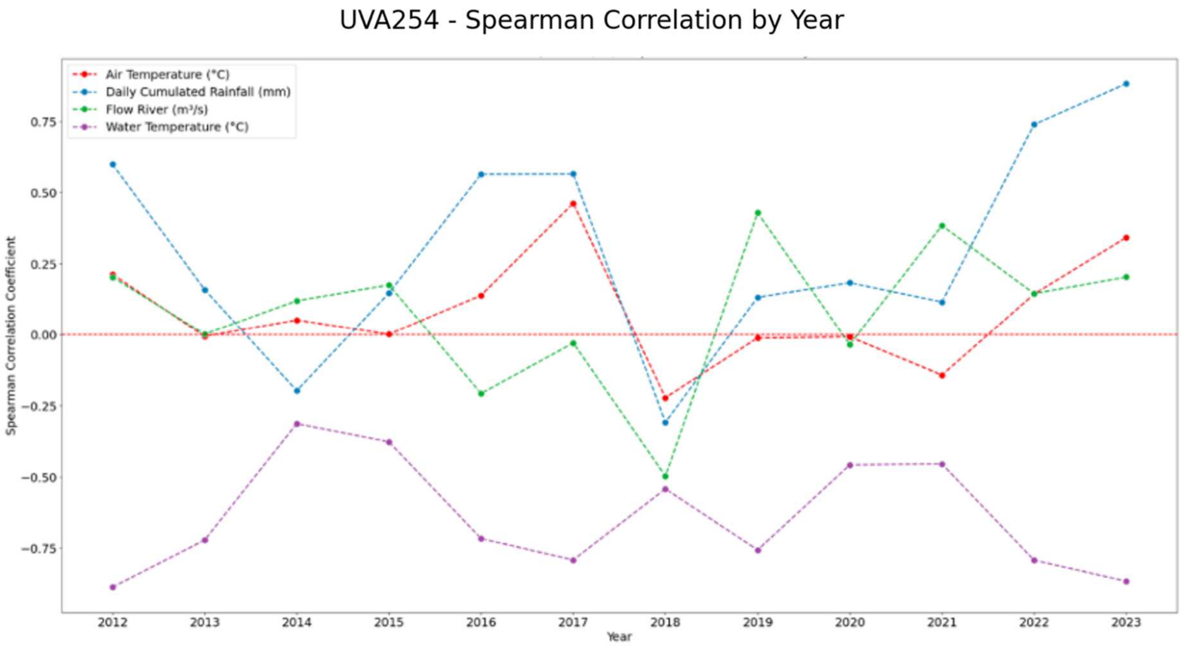

Figure 14. Correlation over the years between the absorbance at 254nm and the other available variables.

The differences between each group of absorbance at 254nm has been studied based on the distribution quartiles of each measured variable. Statistical results were achieved for three out of four variables: air temperature, river flow, and water temperature. Regarding the river flow, statistical evidence that lower values of river flow correspond to lower values of absorbance and vice versa were detected (T-test p-value: 5.86e-08). Conversely, no strong statistical evidence was found for the correspondence of absorbance and rainfall, meaning that based on available data, no statistical relation was found with respect to the absorbance at 254nm. As for the water temperature, it was obtained statistical evidence that lower values of air temperature correspond to higher values of Absorbance at 254nm and vice versa (T-test p-value: 0.00509). Finally, there is statistical evidence that lower values of water temperature correspond to higher values of Absorbance at 254nm and vice versa (T-test p-value: 0.00030). The results have been presented on the figures in Annex 2.

4. Climate change scenarios

4.1 Projections available

Climate projection data come from the Copernicus project (Copernicus EU, 2024). Copernicus is a European Commission program that monitors the Earth and its environment using satellite data and on-site data from sensors. Information is open and free.

Another tool from Copernicus is climate projections, which are obtained by running numerical models for different environmental scenarios. The two studied scenarios were RCP4.5 and RCP8.5: moderated greenhouse gas emissions to the atmosphere, and the high-emission scenario, respectively. The projections were chosen based on the closest coordinates to Xerta and focusing on air and water temperature, rainfall and river flow. Data obtained from the simulations were 30-year averages by month s from 2040 to 2100. Thus, monthly averages from 2041 to 2070, and from 2071 to 2100

4.2 Materials and Methods

In both RCP4.5 and RCP8.5 scenarios, to present the 30-year averaged value, monthly projections are available for three specific years (i.e., 2040, 2070 and 2100). Moreover, monthly historical data are available on Copernicus for year 2000. In order to create a time series for the projections in the entire

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

period 2000-2100, a linear interpolation was performed for each month based on the available projections' years. This operation resulted in a time series with a time step of 1 month for each variable, useful to predict the absorbance 254nm time series until the year 2100.

Trends for the variables available (air and water temperature, rainfall and river flow) from the Copernicus projections and the Tarragona CS dataset were compared. The trend was computed for each variable through the "moving average method" (Johnston et al., 1999). After that, a linear regression model was fitted on each trend to better understand the degree of increase or decrease by looking at the slopes of the fitted models and to assess consistency between the projections' trends and the historical data variables' trends.

Since one of the goals of this task is to predict how the absorbance 254nm will evolve over time based on other water quality parameters' evolutions, regression analysis and modelling were chosen to estimate the target variable (Aasorbance at 254nm (1/m)) with respect to a set of input variables (air temperature (°C), river flow rate (m^3/s), monthly cumulated rainfall (mm), water temperature (°C)). In addition to the reported input variables, different time variables were incorporated into the dataset, such as year, month, season, and timestamp, to let the models capture the time dependency of the time series.

Different Machine Learning (ML) models applied to the Tarragona CS dataset, identified the relationships between the input and the target variables. The model that best approximated the true relationships over a test set was then used to predict the absorbance at 254nm projections based on the input variables. The ML models used were:

- XGBoost (Chen & Guestrin, 2016), an ensemble learning method that primarily employs generalized linear base learners (if the chosen booster is linear). The linear boosting technique sequentially fits simple linear models to the residuals of the previous models, gradually improving the overall prediction accuracy.
- Light GBM (Ke et al., s. f.), a tree-based gradient boosting framework that builds decision trees in a leaf-wise manner, prioritizing the splits that lead to the largest reduction in loss. Linear trees were employed as a strategy to mitigate the extrapolation problem commonly associated with decision trees. This way, the model better handles scenarios where the relationships between input variables and the target variable extend beyond the observed data range.
- Multi-Layer Perceptron (Popescu et al., 2009), an artificial neural network that consists of multiple layers of interconnected neurons, with each layer processing and transferring information to the next one.

To compare the performances of each model, a machine learning pipeline was defined, that is, a set of steps aimed at training, tuning, and testing a machine learning model.

As a first step, the training and the test sets were defined by splitting the dataset with data before and after 2021, with sizes of ~75% and ~25% of the total number of samples, respectively. Furthermore, to balance the impact of each variable and to ensure the correct training of each model, every variable was standardized by removing the mean and scaling to unit variance as follows:

$$x_{scaled} = \frac{x-\mu}{\sigma},$$

where $x \in X$ is a sample of the variable X, μ is the empirical mean and σ the empirical standard deviation of all the samples of X.

Every model was trained following a hyperparameter tuning step, in which the model was optimized by finding the best values for each of its hyperparameters to minimize the Root Mean Squared Error metric (RMSE), defined as the RMSD between a set of predicted and true values. More specifically, the Optuna framework (Akiba et al., 2019) was used to standardize

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

the hyperparameter tuning step, an open-source hyperparameter optimization framework with state-of-the-art algorithms to efficiently search large spaces and prune unpromising trials, leading to faster convergence of the search process. In each trial, every hyperparameter took on a value in the interval defined for that. The performance of a trial was tested by performing time series *k*-fold cross-validation on the training set with *k* = 5 to have more reliable results.

In time series *k*-fold cross-validation, the training set was split into *k* smaller sets. For each split, the test set comprises data points with datetime values higher than those in the corresponding training set. For each of the *k* "folds", the following procedure was then followed:

- a model is trained using *k* - 1 of the folds as training data;
- the resulting model is validated on the remaining part of the data (i.e., it is used as a test set to compute the RMSE).

The performance measure reported by time series *k*-fold cross-validation was calculated as the average of the values computed in the loop, which is the final performance of a specific trial.

For each model, the hyperparameters which were tuned are the following:

- XGBoost:
  - N° Estimators: how many estimators are built within the ensemble model (N° boosting rounds);
  - Eta: Step size shrinkage used in update after every boosting step to prevents overfitting (learning rate);
  - Alpha: L1 regularization term on weights;
  - Lambda: L2 regularization term on weights;
  - Updater: choice of the algorithm to fit linear model.
- Light GBM:
  - N° Estimators: number of decision trees;
  - Learning Rate: shrinking rate;
  - Max Depth: max depth for each tree model;
  - N° Leaves: max number of leaves in one tree;
  - Min Data in Leaf: minimal number of data in one leaf;
  - Lambda L1: L1 regularization term;
  - Lambda L2: L2 regularization term;
  - Min Split Gain: the minimal gain to perform split;
  - Feature Fraction: to randomly select a subset (in %) of features for each tree if < 0 (with resampling);
  - Max Bin: max number of bins that feature values will be bucketed in.
- MLP Neural Network:
  - N° Layers: number of hidden layers of the NN ([1, 2]);
  - N° neurons 1st layer: number of neurons for the first hidden layer [1, 100];
  - N° neurons 2nd layer: number of neurons for the second hidden layer [1, 100];
  - Activation: type of activation function for each hidden layer [relu, identity, logistic, tanh];
  - Solver: type of solver for weight optimization [sgd, adam];
  - Alpha: L2 regularization term;
  - Learning rate: learning rate schedule for weight updates [constant, invsaling, adaptive];
  - Learning rate init: initial learning rate used;
  - Power t: exponent for inverse scaling learning rate (when learning rate is invscaling and solver is sgd);
  - Beta 1: exponential decay rate for estimates of first moment vector in adam;
  - Beta 2: exponential decay rate for estimates of second-moment vector in adam;
  - Epsilon: value for numerical stability in adam.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

Once the best hyperparameters were found, every model was evaluated on the test dataset.

Three evaluation metrics were used to compare the performances of each model. Formally, let

$Y_{test} = \{y_1, \dots, y_n\}$ be the set of the true target values of the test dataset and $\hat{Y}^M_{test} = \{\hat{y}_1, \dots, \hat{y}_n\}$ the set of predicted target values by a generic model M.

The evaluation metrics used were the following:

- Root Mean Squared Error (RMSE): defined as the previously mentioned RMSD where the difference is computed between the predicted values $\hat{y}_i \in \hat{Y}^M_{test}, i = 1, \dots, n$ and the true values $\hat{y}_i \in Y_{test}, i = 1, \dots, n$.

- R^2: coefficient of determination, represents the proportion of the variance in the target variable explained by the predictors. It is defined as

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2},$$

where $\bar{y}$ is the mean of Y_{test}.

- Mean Absolute Error (MAE): measures the average absolute difference between the predicted and the true values, defined as

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|.$$

For each predicted point, a prediction interval was computed to estimate an interval in which the predicted point will fall given what has already been seen. The interval, with a 95% confidence, was computed by using the standard Jackknife Method (Efron, 1992), which is based on a set of leave-one-out models. More specifically:

- for each instance $i = 1, \dots, k$ of the training set, the regression function $\hat{\mu}_{-i}$ is fit on the entire training set with the i^{th} point removed, resulting in *k* leave-one-out models.
- the corresponding leave-one-out conformity score is computed for each i^{th} point as $|y_i - \hat{\mu}_{-i}(x_i)|$.
- the regression function $\hat{\mu}$ is fit on the entire training set and the prediction interval is computed using the computed leave-one-out conformity scores:
$$\hat{\mu}(x_{k+1}) \pm ((1 - \alpha)\ quantile\ of\ |y_1 - \hat{\mu}_{-1}(x_1)|, \dots, |y_k - \hat{\mu}_{-k}(x_k)|.$$
- the resulting confidence interval can, therefore, be summarized as follows:
$C^{jackknife}_{k,\alpha}(x_{k+1}) = [\hat{q}^-_{k,\alpha}\{\hat{\mu}(x_{k+1}) - R^{LOO}_i\}, \hat{q}^+_{k,\alpha}\{\hat{\mu}(x_{k+1}) + R^{LOO}_i\}],$
where $R^{LOO}_i = |y_i - \hat{\mu}_{-i}(x_i)|$.

4.3 Results

The first step was the identification of the best ML model to be used for the projection phase. Each optimized model was trained on 2012-2021 set of data. Then models were tested on the last two years of the UVA254nm. Figure 15 shows the regression predictions for each model, together with a 95% prediction interval, and Figure 16 depicts the zoom of the test period shown in Figure 15.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

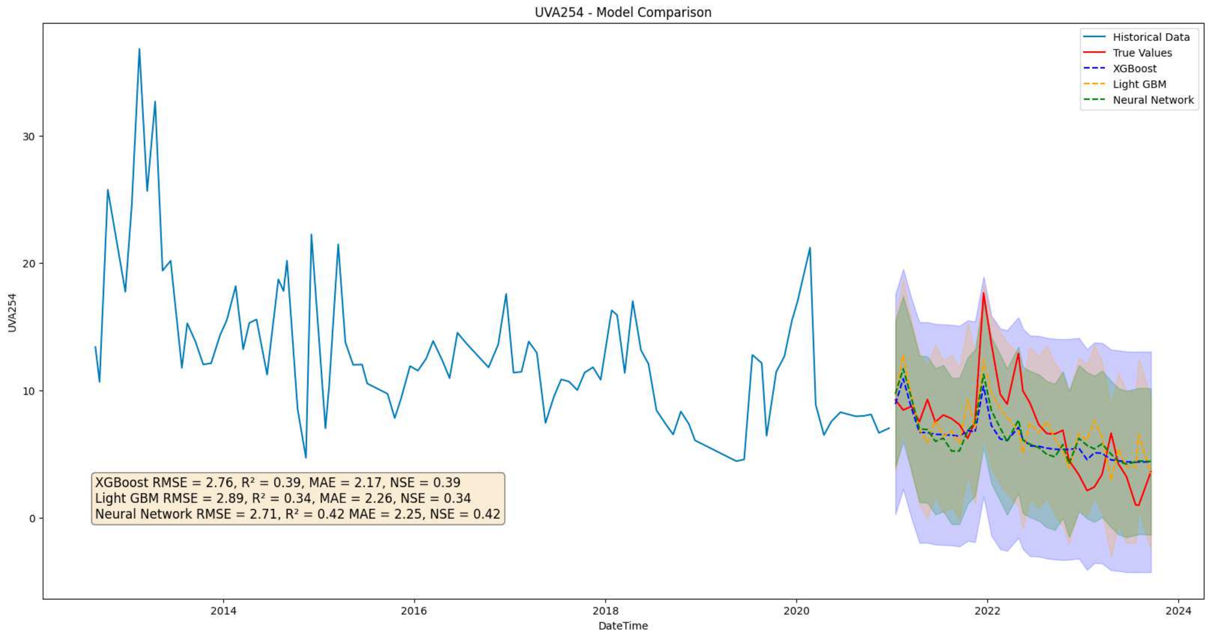

Figure 15. Absorbance at 254nm trend over time. From year 2012 to 2021 corresponds to the training set, while 2021 to 2023 is the test set in which the measured value is compared to the estimated ones according to the different models.

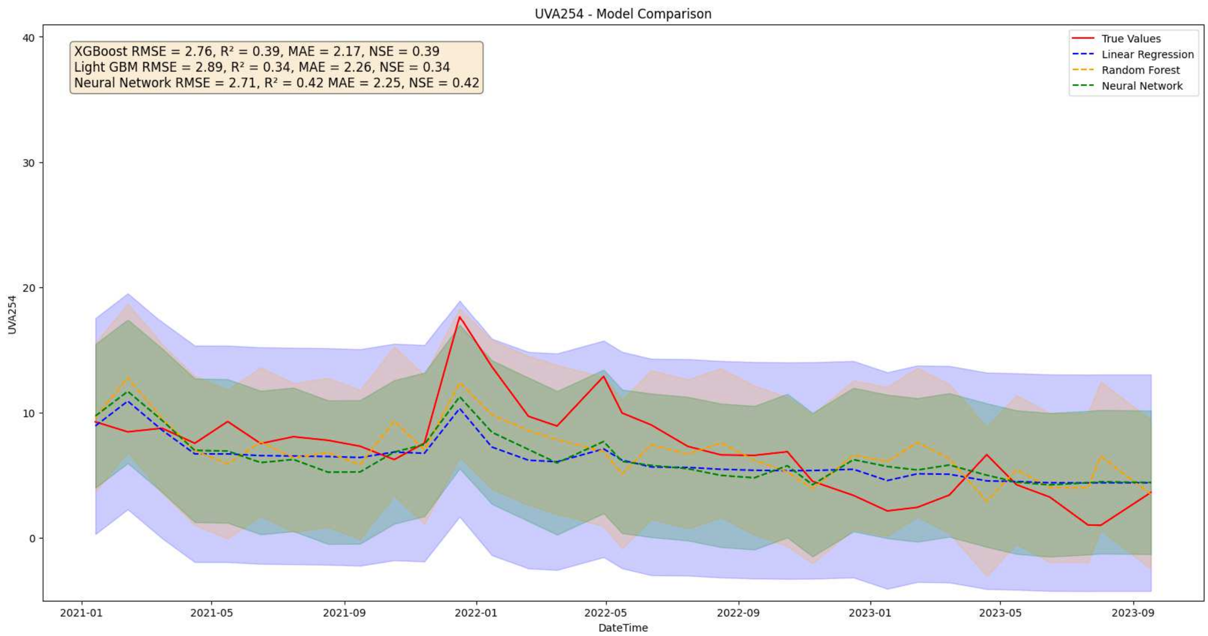

Figure 16. Zoom of the absorbance at 254nm over time for the test period.

Every model was able to catch the decreasing trend of the absorbance at 254nm. Both XGBoost and MLP were less noisy than LightGBM in prediction. The fact that XGBoost and MLP were less noisy was a promising characteristic that might be positive also for the further step of applying them to projections. The MLP NN had a tighter prediction interval than XGBoost, meaning that it was more confident about the prediction of a given value. Based on these considerations and the performance metrics results, MLP was chosen as the best model to be used for projecting the absorbance at 254nm. Table 3 shows the numeric value of the metrics resulting from each model.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

Table 3. Performance metrics for each tested model over the year interval 2021-2023. The best value for each metric is reposted in bold.

	Metrics		
Model	RMSE	R^2	MAE
XGBoost	2.76	0.39	**2.17**
Light GBM	2.89	0.34	2.26
MLP NN	**2.71**	**0.42**	2.25

Having chosen the best prediction model, the second step is the evaluation of future projections. Firstly, the projections trend reported by Copernicus for the input variable were compared to the Xerta measurements available for each input variable (see from Figure 17 to Figure 20). Since the two projections scenarios, namely RCP4.5 and RCP8.5, gave similar results, graphs only related to RCP8.5 are reported.

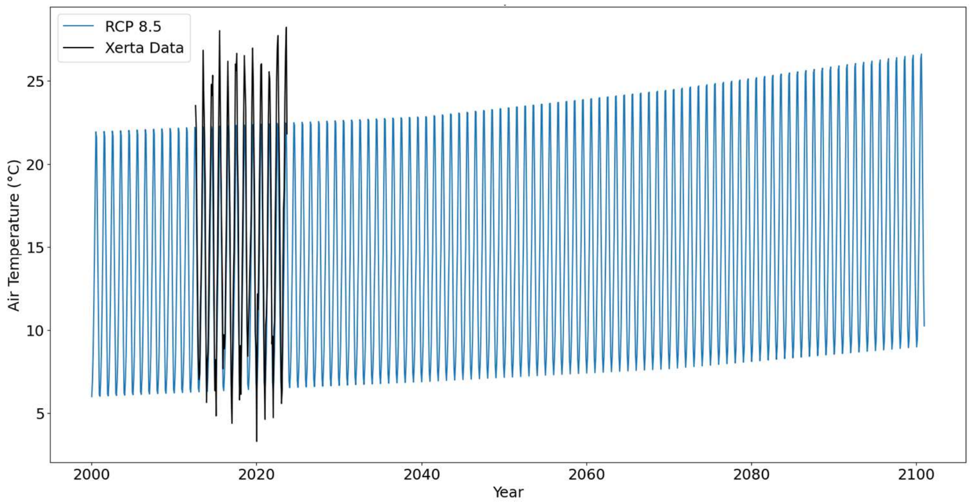

Figure 17. Air Temperature RCP8.5 projections compared to Xerta measurements.

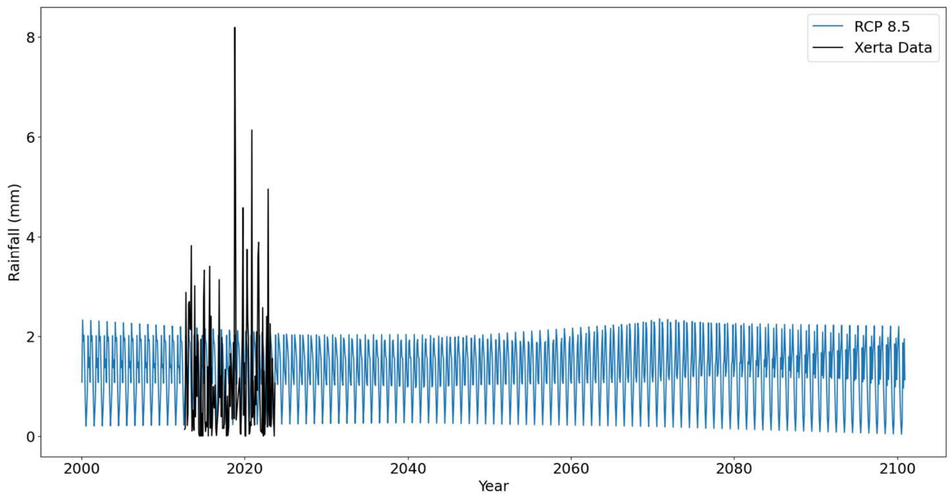

Figure 18. Rainfall RCP8.5 projections compared to Xerta measurements.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

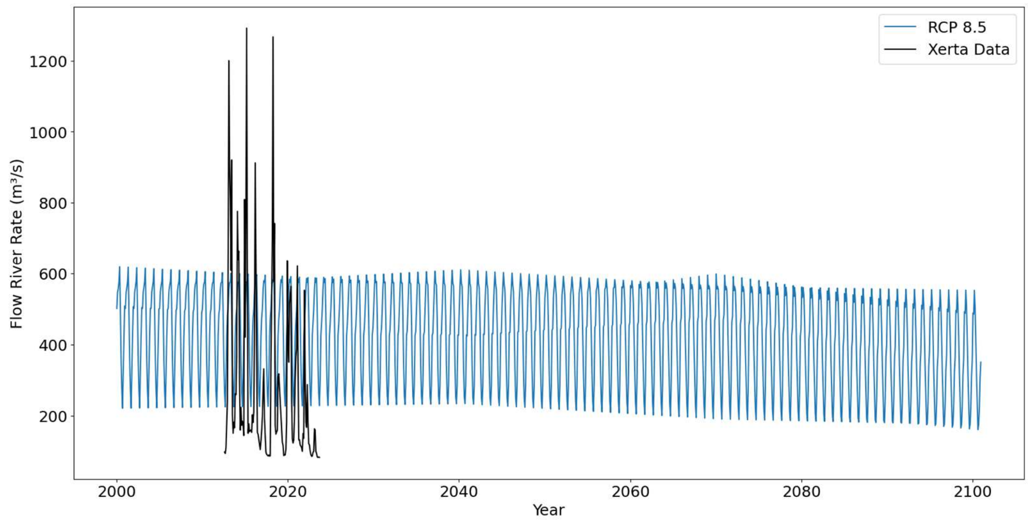

Figure 19. River flow RCP8.5 projections compared to Xerta measurements.

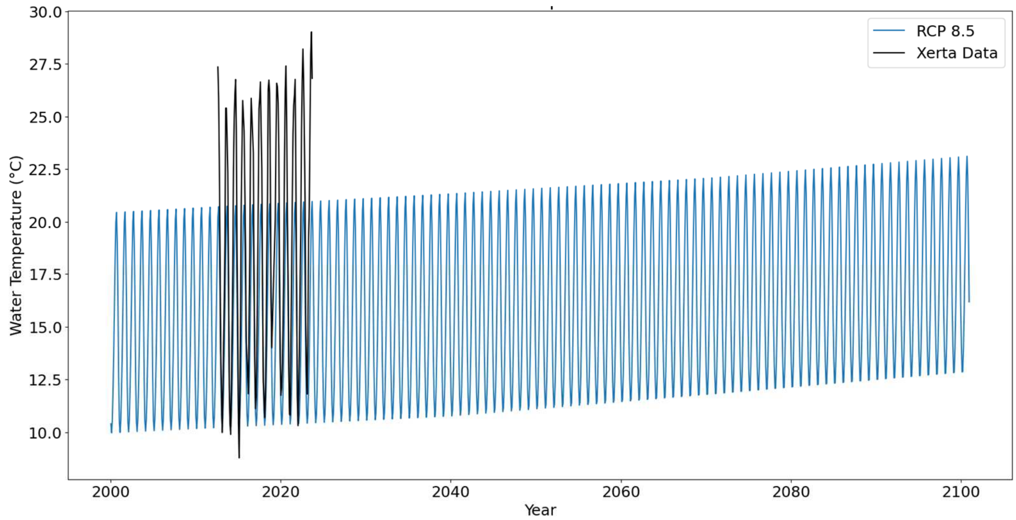

Figure 20. Water Temperature RCP8.5 projections compared to Xerta measurements.

The plots show consistent trends between the projections and the available measurements, especially for the Air Temperature, River flow, and Water Temperature. A common pattern found was that the reconstructed projections tend to underestimate the corresponding variables (Air Temperature, Rainfall and Water Temperature). Regarding the River flow, the projections tend to underestimate (in the warmer months) and overestimate (in the colder months) the Xerta measurements. This behaviour is expected since the projections were reconstructed based on only four years (2010, 2040, 2070, 2100) spread across one hundred years.

Figure 21 and Figure 22 show the projections of absorbance at 254nm performed by the MLP NN model for the RCP4.5 and RCP8.5 scenarios, respectively. For simplicity, the projections for the months in which the UV254 assumes maximum values (march) as well as minimum values (September) are presented in Figure 21. As expected, the projections reveal a decreasing trend in both scenarios, with RCP8.5 slightly decreasing. This is explained by the strong correlation between River flow and absorbance 254nm and the correlation with time (decreasing trend in historical data).

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

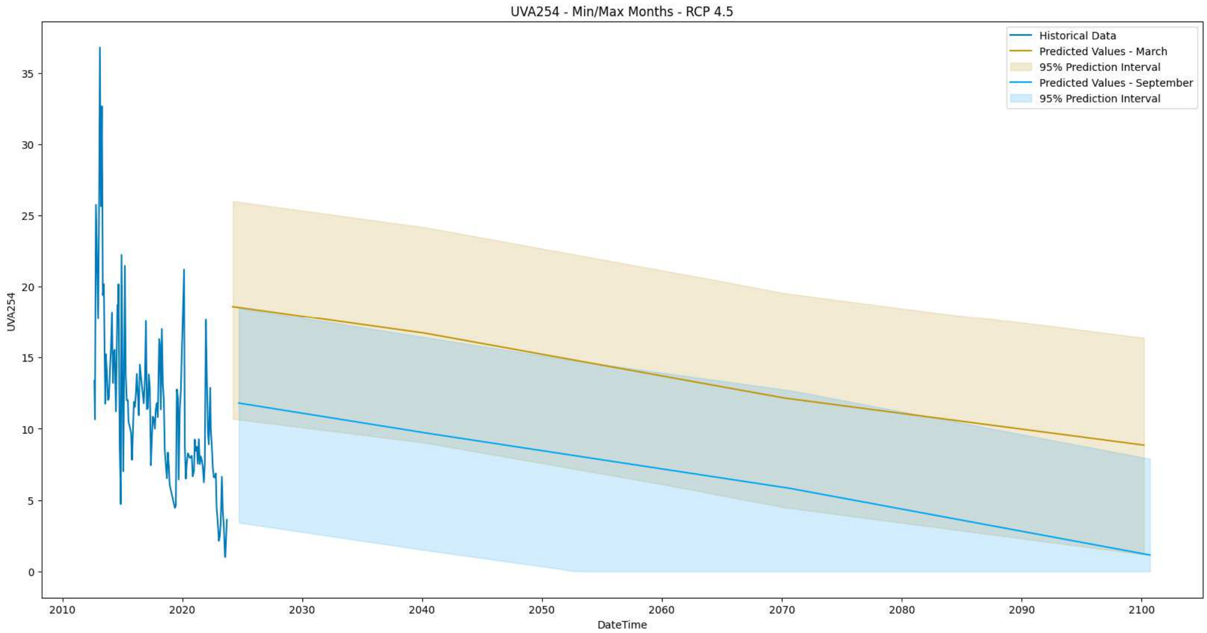

Figure 21. UVA254 projections (RCP4.5).

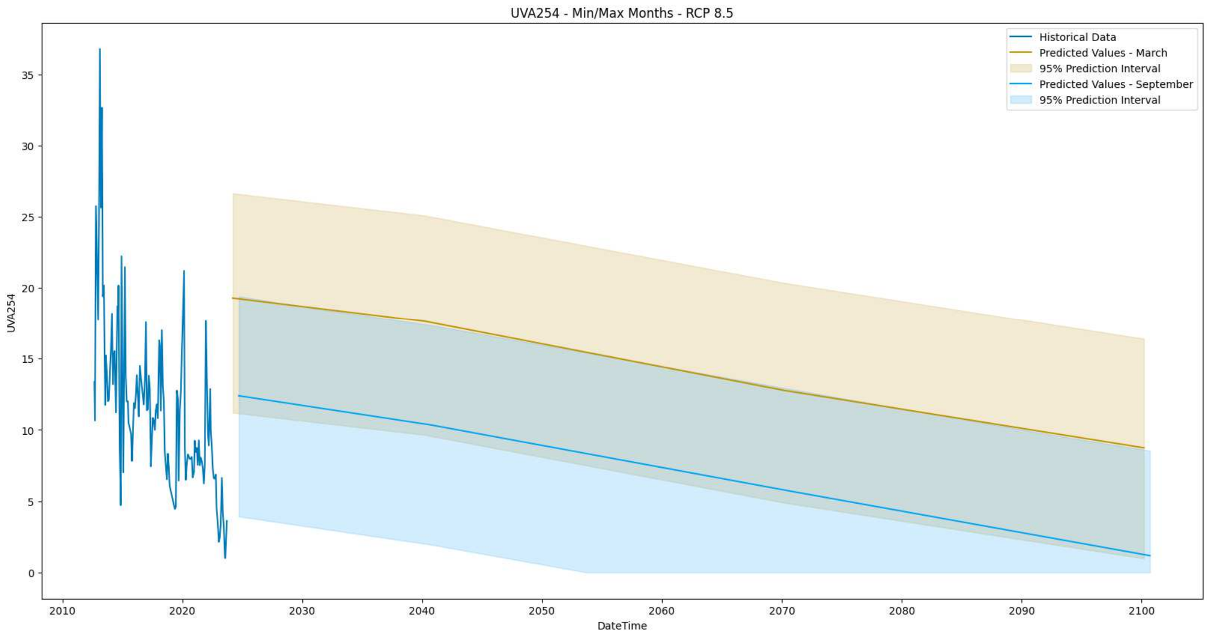

Figure 22. UVA254 projections (RCP8.5).

5. Peak values analysis

Absorbance at 254nm historical data show peaks or extreme values (Figure 23 and Figure 24). These episodes can be complex to manage to drinking water treatment plants due to the possible higher NOM load in water. The aim of this section is to study the relation between these high UVA254 values and the rest of variables (rainfall, water and environment temperature and river flow) to determine if a specific range of variables values can explain the peaks of absorbance.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

5.1 Materials and Methods

Absorbance at 254 peaks were defined by using the Inter Quartile Range (IQR) method for every year taken individually, defining an upper threshold to classify each data point as normal or outlier.

The IQR is the difference between the 25th percentile (Q1) and the 75th percentile (Q3) in a dataset. It measures the spread of the middle 50% of values. Through the IQR method, an observation is declared as peak if it has a value 1.5 times larger than the IQR or 1.5 times less than the IQR.

5.2 Results

As previously described, an upper threshold was defined for every year to classify when an observation is an outlier and when it is considered normal. Figure 23 shows the upper IQR thresholds for each year of the absorbance at 254nm (every year with a different color for visual clarity). The values above each threshold were considered outliers, as can be seen in Figure 24, where the red dots represent the observations marked as outliers.

A noteworthy anomaly becomes visible at the beginning of 2020, where a great number of values were classified as outliers. This is due to a combination of the data reconstruction phase and to an extreme event happened at the end of January 2020, storm Gloria, which primarily affected eastern Spain. The correspondence of an anomalous peak was found also in the timeseries of Rainfall.

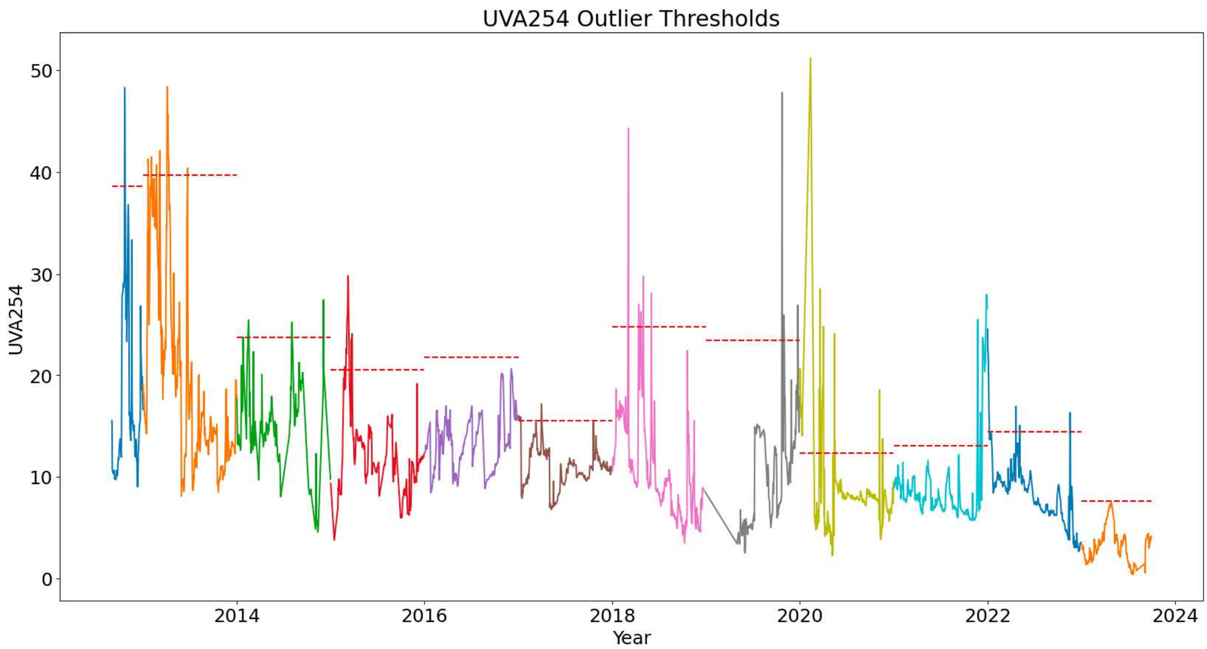

Figure 23. Absorbance trend over time with reported upper absorbance thresholds for every year calculated through IQR method.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

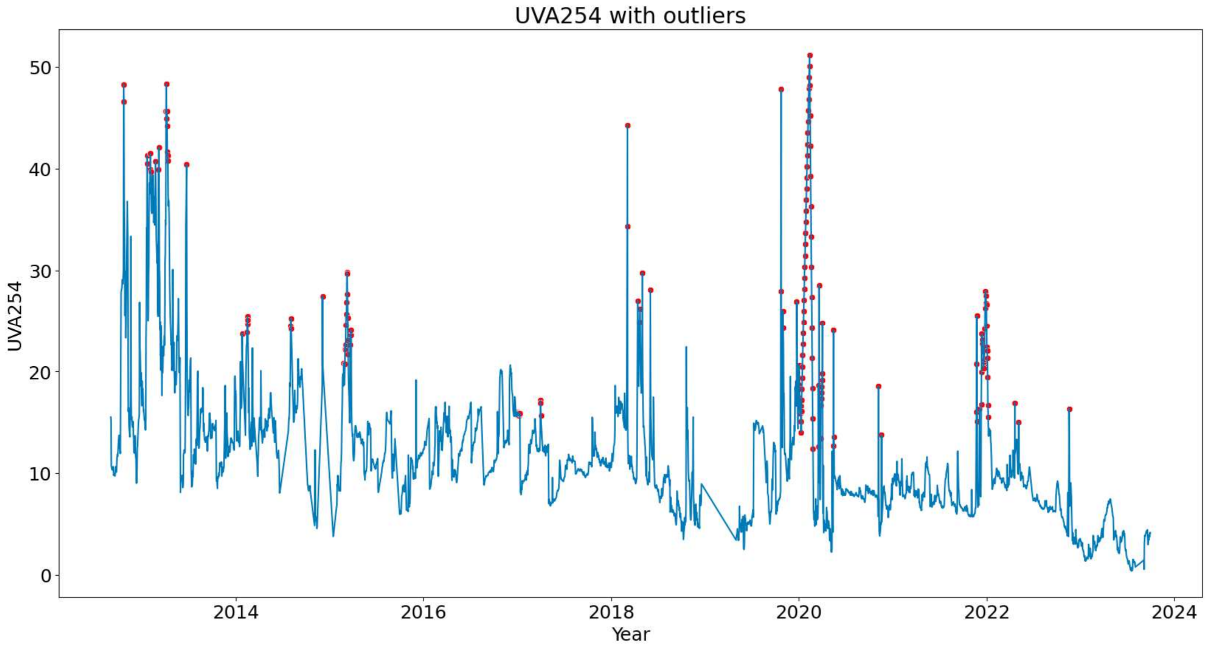

Figure 24. Absorbance trend over time with reported observations marked as peaks (red dots).

From Figure 25Figure 25 to Figure 28Figure 28, they present the differences between the values of each variable corresponding to normal and outlier values of Absorbance at 254nm. Consistency is observed regarding the fact that peak values of absorbance at 254nm happen mostly during colder months as the distributions of the air and water temperatures are lower for the outliers compared to the normal values distributions (see Figure 25 and Figure 28, respectively). Also, uniformity was found with previous analyses regarding the river flow in terms of high values of Absorbance at 254nm corresponding to high values of river flow (see Figure 27Figure 27). In lower Ebro area (after Zaragosa, as this case study), high river flow is related to dam's gate opening, but not with rainfall (rainfall is significantly lower in Catalonia compared to upper regions). In winter, Mequinenza dam normally needs to make room to absorb water from the upper part of the river or other river effluents caused by the meltdown. This is coherent with data: positive correlation between low temperature (winter), high flow river due to dam's gate opening, and higher UVA254 values (i.e. turbidity raise due to sediments and NOM addition to the river).

Regarding the rainfall, there is a slight difference between the two probability density functions depicted in Figure 26 as well as between the IQRs of the two boxplots presented in Annex 3 (where we show just the values in (0, 30) for visual clarity). However, there is not a clear difference that could represent a dependency between high Rainfall days and peak absorbance at 254nm days, in fact the boxplots of Annex 3 related to rainfall show that there are a lot of high values of Rainfall corresponding to absorbance at 254nm values classified as normal. In addition, Xerta rainfall is not uniquely the one affecting to Ebro river flow, but mainly the rains in upstream regions. So, only focusing on the closer station to the water intake is not representative to study the rainfall impact on the river.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

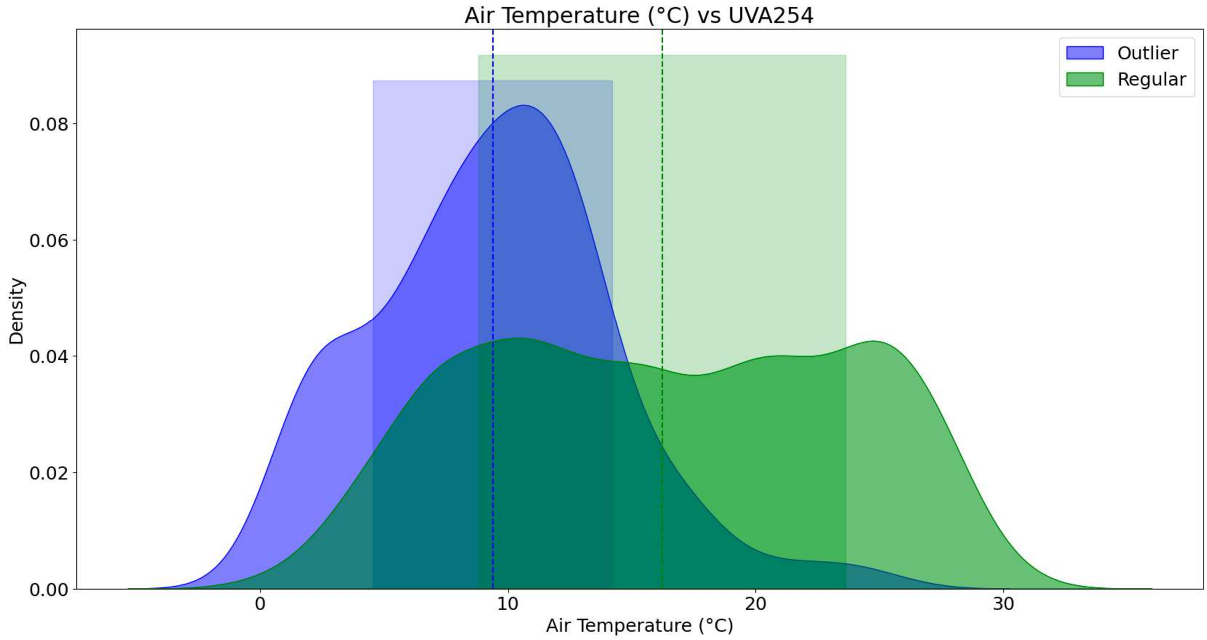

Figure 25. Air temperature density plot corresponding to absorbance at 254nm normal and outlier values.

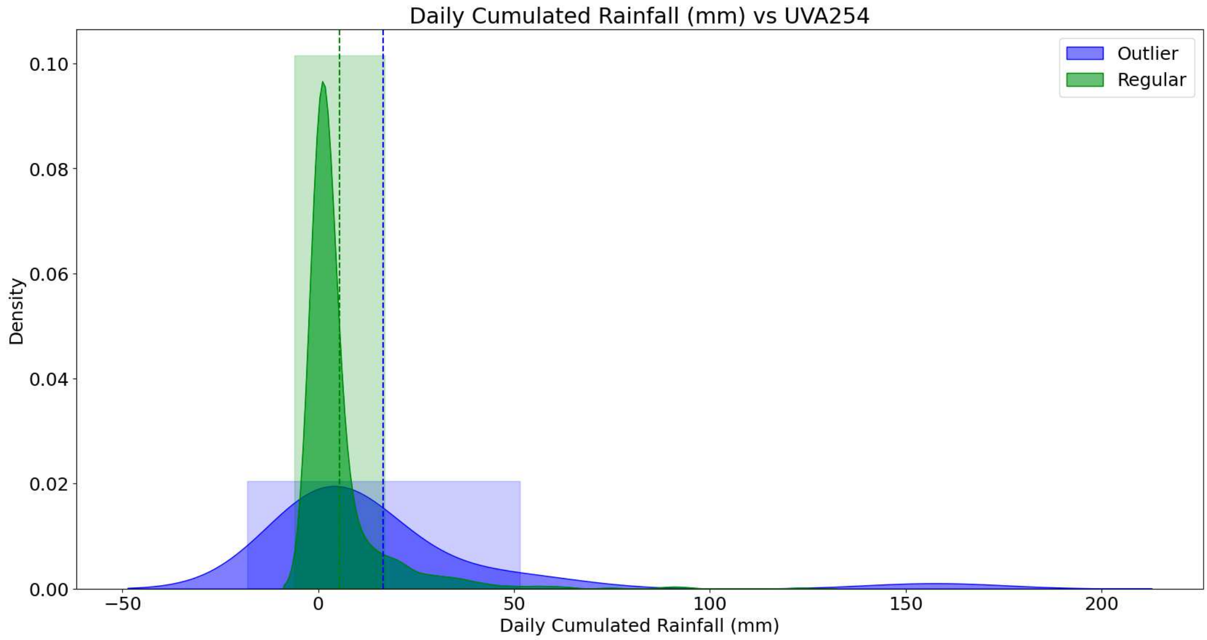

Figure 26. Rainfall density plot corresponding to absorbance at 254nm normal and outlier values (only values > 0).

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

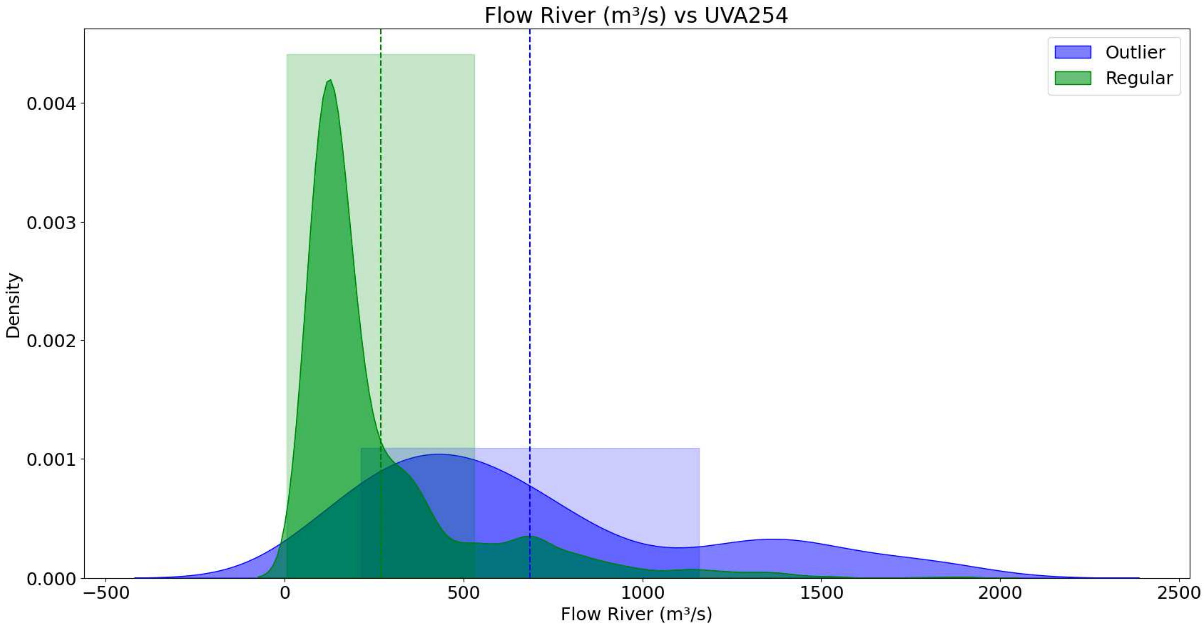

Figure 27. River flow density plot corresponding to absorbance at 254nm normal and outlier values.

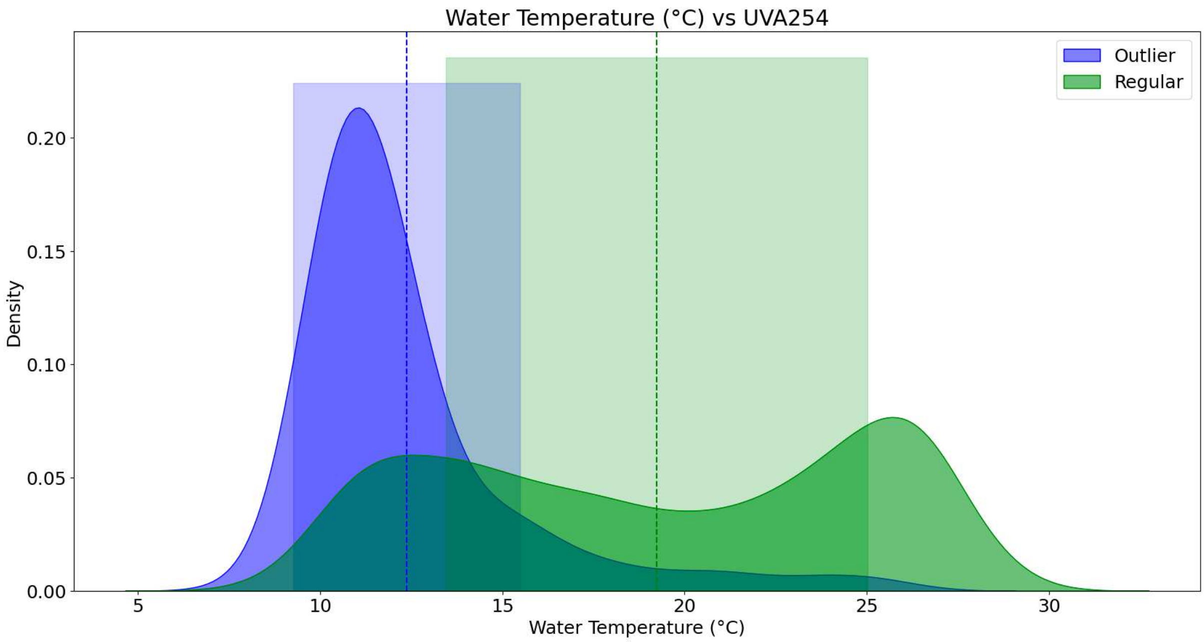

Figure 28. Water Temperature boxplots corresponding to absorbance at 254nm normal and outlier values.

6. Discussion

Based on the state of the art performed, the climatic drivers that can impact the quantity and the quality of water sources used to produce drinking water were clearly identified: precipitation and air temperature. These variables have been recorded for dozens of years, which have allowed to study the historic and current observed trends. Additionally, the large datasets available promotes the implementation and validation of climate models at regional and global scales for future projections. Thus, for these climate drivers, quantitative projections at different horizons and for different scenarios exist and are commonly disseminated (e.g., projections presented by Copernicus).

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

The impact of these climatic drivers on water sources has been less studied due to the lack of available information and knowledge. Therefore, the existing literature mainly presents the potential interlinks between the climatic drivers and the water quantity and quality. Projections on surface water quantity and water temperature are also available, though there is still an uncertainty on how good these estimations are. Despite the clear evidence on their impact on water sources, there are not quantitative projections of other relevant water quality variables that condition the DWSS management and adaptation. Furthermore, water quality projections should be downscaled at regional level considering watershed management conditions, local economic activities, regional regulations, etc.

Tarragona case study (CS#3) is in the Mediterranean region of the North-East of Spain, where high temperatures and longer drought periods will become more frequent due to climate change. Historical monitoring data allowed to study these trends, and a large data set permit to build and implement data-driven models to evaluate the progress of the studied system. Hence, this modelling allows to predict future trends and advancing to possible extreme events.

Historical data from different stations at different locations show that river flow is decreasing, and temperatures and conductivity are raising. This new context could affect distribution network: higher water temperatures affect directly disinfection processes, and the kinetic of DBPs production. Chemical disinfectants are affected through fasting degradation (i.e. ozone) or volatilization (i.e. chlorine). In addition, pathogens grow up easier at higher temperatures, so disinfection might have to be intensified to ensure microbial safe water. Finally, adding more chemical disinfectant, in general will also come up with higher DBP formation in a synergic way: more disinfectant and more formation ratio.

River flow decrease will affect water availability and water quality (salts, nutrients and organic matter will be more concentrated). Water scarcity can affect industrial and agronomics sectors, and consumers habits. Water use has been already regulated in some regions of Catalonia to decrease water consumption depending on the water availability. As a consequence, water reuse has grown for non-potable uses to cover the water demand at the different economic sectors and, hence, minimising stress on the water systems.

Conductivity is also related to river flow: salts will be more concentrated, leading to rising conductivity values. DBPs speciation can be affected by bromine or iodine concentration increase: other halogenated DBPs than chlorine-DBPs might boost.

Contrary to most literature, organic matter (either NOM or here represented as UV-absorbance at 254 terms) trend is decreasing for this case study analysis as well as the ammonium concentration. This might be the result of more restrictive wastewater regulations: industries and municipalities must discharge a higher quality water effluent than historically (see Figure 29). Thus, in general terms, NOM increase might not be a problem in case of Tarragona and Ebro River, but water availability will be. Government and water entities will have to cooperate to ensure secure and safe water access to population considering alternatives water sources (regenerated and desalinated water) and water consumptions limitations if needed. Surface water will not be the unique source anymore in frequent drought regions as the Mediterranean.

UVA245 concentration, here assumed as NOM, along the last decade is decreasing but shows anomalies: peaks of absorbance not related with rainfall. These discrete episodes can be explained by the basin management based on dams (see Figure 29). Thus, the development and implementation of an early warning system to warn water utilities when a NOM peak will reach the DWTP intake would be useful to adapt the treatment when needed. In this way, digital tools for water management and predicting when dams will be opened to discharge higher water volume from the reservoirs, added to the prediction of river flows or direct information from the basin management entities, would be useful to adjust the drinking water treatment to ensure its quality.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

Future work should be connected to the study of a potential NOM composition change due to climate impacts. This task focused on NOM amount but is well known that DBPs production is linked with some fractions of NOM (e.g., aromatic bonds), so NOM composition. Studying which fractions of NOM may rise, and which may diminish due to climate change will give us a deeper insight into which DBP precursors would reach the intake points of DWTP, and if the DWTP are not designed to totally remove these precursors will reach out the outlet of the plant.

Land uses annexed to the watershed might also impact the water quality either having a positive or a negative impact (see Figure 29). Thus, including the land uses into the watershed management model can contribute to the prediction of the water quality.

Finally, despite the historical data show clear trends for the last 10 years in Tarragona, a further exploration of the extreme events (such as prolonged droughts and severe storms) and their consequences should be performed in order to identify the impact and variability of water quantity and quality. Hence, this knowledge can support the improvement of DWTP exploitation plans and the adaptation the DWSS to increase their versatility and flexibility in front of extreme events due to climate change.

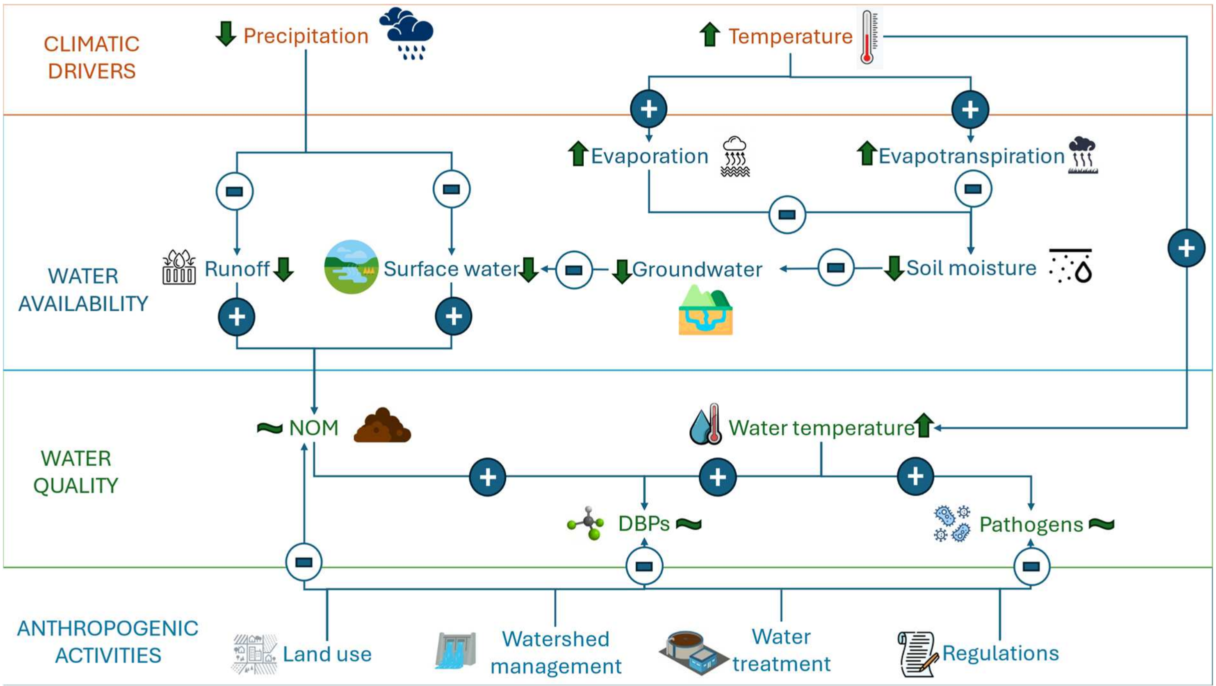

Figure 29. Summary of the impacts on water quality including anthropogenic activities.

7. Conclusions and tipping points

Climate change is already affecting water bodies both in terms of quantity and quality. At the Mediterranean area, cycles are getting more extreme: larger drought periods and higher intense rain episodes. For the utilities using surface water, drought periods involve river flow decrease, salts are more concentrated (conductivity rise) and wastewater treatment plant (WWTP) effluent (with a special attention to emerging compounds) gain proportion on the streams and then used as water source for the DWTP. On the other hand, intense rainfall episodes dilute the water arriving to the DWTP intake, improving water quality in general terms. Thus, DWTP using surface water will have to

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

be more versatile and prepared to treat different water qualities sources: able to treat high conductivity water and to remove emerging contaminants from the WWTP, but also high-quality raw water, present in rainy periods or meltdown season.

Legislation will have an important role too: WWTP effluents might be more restrictive to limit emerging contaminants concentrations leaked to the river (water source for utilities using surface water). In addition, future WWTP need to have the capacity to treat more volume, to be prepared for the rainy episodes without collapsing municipalities.

For the utilities without disinfection treatment, temperature rise might boost the growth of pathogens in raw water. Therefore, they might have to disinfect at high temperature periods and start preparing for DBPs precursors removal if needed. Additionally, further studies about how pathogens can degrade DBPs precursors, such as NOM, with an increase of the water temperature and solar radiation should be performed their impact on NOM concentrations.

To sum up, the main variables that need to be controlled are water temperature, conductivity, NOM and pathogens. Future work will include to study the tipping points of these variables, what would be useful for water utilities readiness for climate change impact.

8. Bibliography

Amanambu, A. C., Obarein, O. A., Mossa, J., Li, L., Ayeni, S. S., Balogun, O., Oyebamiji, A., & Ochege, F. U. (2020). Groundwater system and climate change: Present status and future considerations. *Journal of Hydrology*, *589*, 125163. https://doi.org/10.1016/j.jhydrol.2020.125163

Anderson, L. E., DeMont, I., Dunnington, D. D., Bjorndahl, P., Redden, D. J., Brophy, M. J., & Gagnon, G. A. (2023). A review of long-term change in surface water natural organic matter concentration in the northern hemisphere and the implications for drinking water treatment. *Science of The Total Environment*, *858*, 159699. https://doi.org/10.1016/j.scitotenv.2022.159699

Bangash, R. F., Passuello, A., Sanchez-Canales, M., Terrado, M., López, A., Elorza, F. J., Ziv, G., Acuña, V., & Schuhmacher, M. (2013). Ecosystem services in Mediterranean river basin: Climate change impact on water provisioning and erosion control. *Science of The Total Environment*, *458-460*, 246-255. https://doi.org/10.1016/j.scitotenv.2013.04.025

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

Blaszczak, J. R. (2023). Deoxygenation of temperate rivers. *Nature Climate Change*, *13*(10), 1021-1022.

CAT. (2024). *Consorci d'Aigües de Tarragona*. https://www.ccaait.com/

CHE. (2024). *Confederación Hidrográfica del Ebro*. https://www.chebro.es/web/guest/la-confederacion

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785-794. https://doi.org/10.1145/2939672.2939785

Clark, J. M., Chapman, P. J., Adamson, J. K., & Lane, S. N. (2005). Influence of drought-induced acidification on the mobility of dissolved organic carbon in peat soils. *Global Change Biology*, *11*(5), 791-809. https://doi.org/10.1111/j.1365-2486.2005.00937.x

Clark, J. M., Lane, S. N., Chapman, P. J., & Adamson, J. K. (2007). Export of dissolved organic carbon from an upland peatland during storm events: Implications for flux estimates. *Journal of Hydrology*, *347*(3-4), 438-447. https://doi.org/10.1016/j.jhydrol.2007.09.030

Cool, G., Delpla, I., Gagnon, P., Lebel, A., Sadiq, R., & Rodriguez, M. J. (2019). Climate change and drinking water quality: Predicting high trihalomethane occurrence in water utilities supplied by surface water. *Environmental Modelling & Software*, *120*, 104479. https://doi.org/10.1016/j.envsoft.2019.07.004

Copernicus EU. (2024). *Copernicus—Earth Observation component of the European Union's Space Programme*. https://www.copernicus.eu/en

Couture, S., Houle, D., & Gagnon, C. (2012). Increases of dissolved organic carbon in temperate and boreal lakes in Quebec, Canada. *Environmental Science and Pollution Research*, *19*(2), 361-371. https://doi.org/10.1007/s11356-011-0565-6

Creed, I. F., Bergström, A.-K., Trick, C. G., Grimm, N. B., Hessen, D. O., Karlsson, J., Kidd, K. A., Kritzberg, E., McKnight, D. M., Freeman, E. C., Senar, O. E., Andersson, A., Ask, J., Berggren, M., Cherif, M., Giesler, R., Hotchkiss, E. R., Kortelainen, P., Palta, M. M., … Weyhenmeyer, G.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

A. (2018). Global change-driven effects on dissolved organic matter composition: Implications for food webs of northern lakes. *Global Change Biology*, *24*(8), 3692-3714. https://doi.org/10.1111/gcb.14129

De Wit, H. A., Valinia, S., Weyhenmeyer, G. A., Futter, M. N., Kortelainen, P., Austnes, K., Hessen, D. O., Räike, A., Laudon, H., & Vuorenmaa, J. (2016). Current Browning of Surface Waters Will Be Further Promoted by Wetter Climate. *Environmental Science & Technology Letters*, *3*(12), 430-435. https://doi.org/10.1021/acs.estlett.6b00396

Delpla, I., Jung, A.-V., Baures, E., Clement, M., & Thomas, O. (2009). Impacts of climate change on surface water quality in relation to drinking water production. *Environment International*, *35*(8), 1225-1233. https://doi.org/10.1016/j.envint.2009.07.001

Delpla, I., & Rodriguez, M. J. (2014). Effects of future climate and land use scenarios on riverine source water quality. *Science of The Total Environment*, *493*, 1014-1024. https://doi.org/10.1016/j.scitotenv.2014.06.087

EEA. (2024). *Water use in Europe—Quantity and quality face big challenges*. Environmental European Agency. https://www.eea.europa.eu/signals-archived/signals-2018-content-list/articles/water-use-in-europe-2014#:~:text=About%2088.2%20%25%20of%20Europe's%20freshwater,exploitation%2C%20pollution%20and%20climate%20change.

Efron, B. (1992). Jackknife-After-Bootstrap Standard Errors and Influence Functions. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *54*(1), 83-111. https://doi.org/10.1111/j.2517-6161.1992.tb01866.x

Evans, C. D. (2005). Modelling the effects of climate change on an acidic upland stream. *Biogeochemistry*, *74*(1), 21-46. https://doi.org/10.1007/s10533-004-0154-6

Evans, C. D., Monteith, D. T., Fowler, D., Cape, J. N., & Brayshaw, S. (2011). Hydrochloric Acid: An Overlooked Driver of Environmental Change. *Environmental Science & Technology*, *45*(5), 1887-1894. https://doi.org/10.1021/es103574u

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

Evans, M., Warburton, J., & Yang, J. (2006). Eroding blanket peat catchments: Global and local implications of upland organic sediment budgets. *Geomorphology*, *79*(1-2), 45-57. https://doi.org/10.1016/j.geomorph.2005.09.015

Finstad, A. G., Andersen, T., Larsen, S., Tominaga, K., Blumentrath, S., De Wit, H. A., Tømmervik, H., & Hessen, D. O. (2016). From greening to browning: Catchment vegetation development and reduced S-deposition promote organic carbon load on decadal time scales in Nordic lakes. *Scientific Reports*, *6*(1), 31944. https://doi.org/10.1038/srep31944

Garnier, M., & Holman, I. (2019). Critical Review of Adaptation Measures to Reduce the Vulnerability of European Drinking Water Resources to the Pressures of Climate Change. *Environmental Management*, *64*(2), 138-153. https://doi.org/10.1007/s00267-019-01184-5

Hofstra, N. (2011). Quantifying the impact of climate change on enteric waterborne pathogen concentrations in surface water. *Current Opinion in Environmental Sustainability*, *3*(6), 471-479. https://doi.org/10.1016/j.cosust.2011.10.006

IPCC. (2014). IPCC fifth assessment report: Impacts, adaptation and vulnerability. *Part A: global and sectoral aspects. Contribution of working group II to the fifth assessment report of the intergovernmental Panel on Climate Change*, *1132*.

IPCC. (2021). *Climate Change 2021: The Physical Science Basis*.

IPPC. (2022). *IPCC sixth assessment report: Impacts, adaptation and vulnerability*.

Johnston, F. R., Boyland, J. E., Meadows, M., & Shale, E. (1999). Some properties of a simple moving average when applied to forecasting a time series. *Journal of the Operational Research Society*, *50*(12), 1267-1271. https://doi.org/10.1057/palgrave.jors.2600823

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (s. f.). *LightGBM: A Highly Efficient Gradient Boosting Decision Tree*.

Kouakou, C. R. C., & Poder, T. G. (2019). Economic impact of harmful algal blooms on human health: A systematic review. *Journal of Water and Health*, *17*(4), 499-516. https://doi.org/10.2166/wh.2019.064

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

Kristensen, P. (1996). *Water quality of large rivers*. European Environment Agency.

Kumar, C. P. (2012). Climate change and its impact on groundwater resources. *International Journal of Engineering and Science*, *1*(5), 43-60.

Leveque, B., Burnet, J.-B., Dorner, S., & Bichai, F. (2021). Impact of climate change on the vulnerability of drinking water intakes in a northern region. *Sustainable Cities and Society*, *66*, 102656. https://doi.org/10.1016/j.scs.2020.102656

Levine, A. D., Yang, Y. J., & Goodrich, J. A. (2016). Enhancing climate adaptation capacity for drinking water treatment facilities. *Journal of Water and Climate Change*, *7*(3), 485-497. https://doi.org/10.2166/wcc.2016.011

Li, Z., Clark, R. M., Buchberger, S. G., & Jeffrey Yang, Y. (2014). Evaluation of Climate Change Impact on Drinking Water Treatment Plant Operation. *Journal of Environmental Engineering*, *140*(9), A4014005. https://doi.org/10.1061/(ASCE)EE.1943-7870.0000824

Lipczynska-Kochany, E. (2018). Effect of climate change on humic substances and associated impacts on the quality of surface water and groundwater: A review. *Science of the total environment*, *640*, 1548-1565.

Lopez-Bustins, J. A., Arbiol-Roca, L., Martin-Vide, J., Barrera-Escoda, A., & Prohom, M. (2019). *Intra-annual variability of the Western Mediterranean Oscillation (WeMO) and occurrence of extreme torrential rainfall in Catalonia (NE Iberia)*. https://doi.org/10.5194/nhess-2019-374

Ma, B., Hu, C., Zhang, J., Ulbricht, M., & Panglisch, S. (2022). Impact of Climate Change on Drinking Water Safety. *ACS ES&T Water*, *2*(2), 259-261. https://doi.org/10.1021/acsestwater.2c00004

Mitchell, M. J., & Likens, G. E. (2011). Watershed Sulfur Biogeochemistry: Shift from Atmospheric Deposition Dominance to Climatic Regulation. *Environmental Science & Technology*, *45*(12), 5267-5271. https://doi.org/10.1021/es200844n

Moore, S. K., Trainer, V. L., Mantua, N. J., Parker, M. S., Laws, E. A., Backer, L. C., & Fleming, L. E. (2008). Impacts of climate variability and future climate change on harmful algal blooms and

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

human health. *Environmental Health*, *7*(Suppl 2), S4. https://doi.org/10.1186/1476-069X-7-S2-S4

Parry, M. L. (2007). *Climate change 2007-impacts, adaptation and vulnerability: Working group II contribution to the fourth assessment report of the IPCC* (Vol. 4). Cambridge University Press.

Popescu, M.-C., Balas, V. E., Perescu-Popescu, L., & Mastorakis, N. (2009). *Multilayer Perceptron and Neural Networks*. *8*(7).

Prathumratana, L., Sthiannopkao, S., & Kim, K. W. (2008). The relationship of climatic and hydrological parameters to surface water quality in the lower Mekong River. *Environment International*, *34*(6), 860-866. https://doi.org/10.1016/j.envint.2007.10.011

Ritson, J. P., Graham, N. J. D., Templeton, M. R., Clark, J. M., Gough, R., & Freeman, C. (2014). The impact of climate change on the treatability of dissolved organic matter (DOM) in upland water supplies: A UK perspective. *Science of The Total Environment*, *473-474*, 714-730. https://doi.org/10.1016/j.scitotenv.2013.12.095

SAIH Ebro. (2024). *Sistema Automático de Información Hidrológica de la Cuenca Hidrográfica del Ebro*. http://www.saihebro.com/saihebro/index.php?url=/autoservicio/inicio

Sardans, J., Peñuelas, J., & Estiarte, M. (2008). Changes in soil enzymes related to C and N cycle and in soil C and N content under prolonged warming and drought in a Mediterranean shrubland. *Applied Soil Ecology*, *39*(2), 223-235. https://doi.org/10.1016/j.apsoil.2007.12.011

Sigg, L. (2000). Redox Potential Measurements in Natural Waters: Significance, Concepts and Problems. En J. Schüring, H. D. Schulz, W. R. Fischer, J. Böttcher, & W. H. M. Duijnisveld (Eds.), *Redox* (pp. 1-12). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-04080-5_1

Strock, K. E., Saros, J. E., Nelson, S. J., Birkel, S. D., Kahl, J. S., & McDowell, W. H. (2016). Extreme weather years drive episodic changes in lake chemistry: Implications for recovery from sulfate deposition and long-term trends in dissolved organic carbon. *Biogeochemistry*, *127*(2-3), 353-365. https://doi.org/10.1007/s10533-016-0185-9

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

UNESCO. (2020). *Water and Climate Change*.

Vineis, P., Chan, Q., & Khan, A. (2011). Climate change impacts on water salinity and health. *Journal of Epidemiology and Global Health*, *1*(1), 5. https://doi.org/10.1016/j.jegh.2011.09.001

Watmough, S. A., Eimers, C., & Baker, S. (2016). Impediments to recovery from acid deposition. *Atmospheric Environment*, *146*, 15-27. https://doi.org/10.1016/j.atmosenv.2016.03.021

Wikipedia. (2024). *Cuenca Hidrográfica del Ebro*. https://es.wikipedia.org/wiki/Cuenca_hidrogr%C3%A1fica_del_Ebro

Wilby, R. L., & Wigley, T. M. (1997). Downscaling general circulation model output: A review of methods and limitations. *Progress in physical geography*, *21*(4), 530-548.

WISE. (2024). *Water Sources of Europe*. Freshwater Information System for Europe. https://water.europa.eu/freshwater/europe-freshwater/freshwater-themes/water-resources-europe

Xiao, Y., Rohrlack, T., & Riise, G. (2020). Unraveling long-term changes in lake color based on optical properties of lake sediment. *Science of The Total Environment*, *699*, 134388. https://doi.org/10.1016/j.scitotenv.2019.134388

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

Annex 1

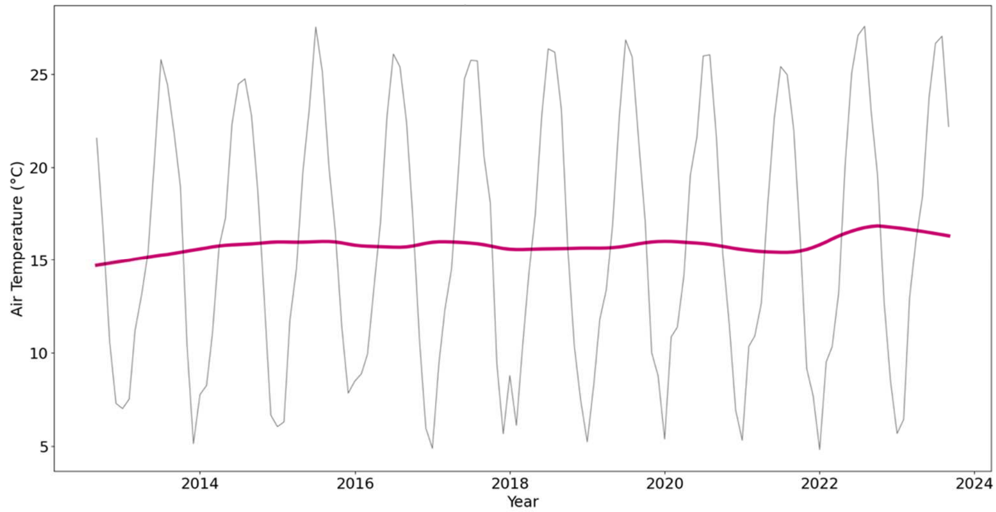

Figure 30. Monthly average time series (black line) and the corresponding moving average (red line) for air temperature.

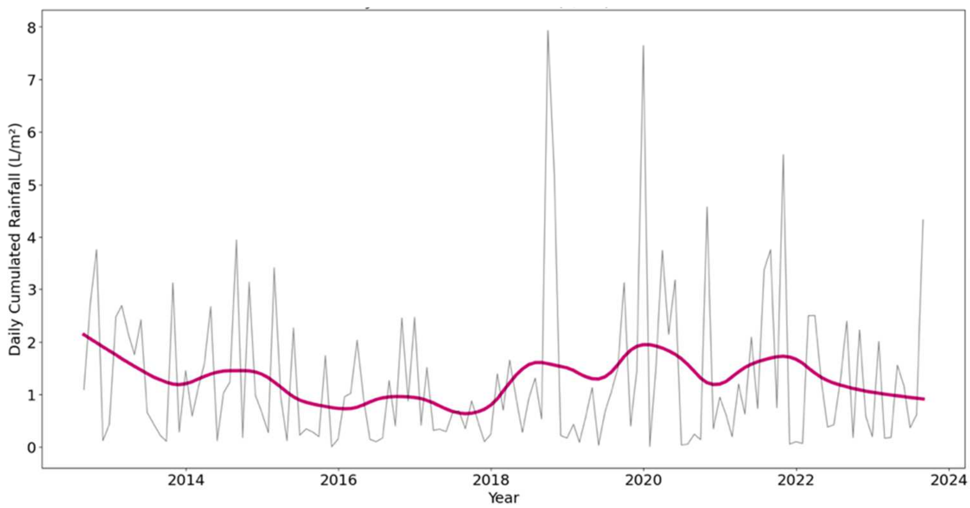

Figure 31. Monthly average time series (black line) and the corresponding moving average (red line) for air temperature.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

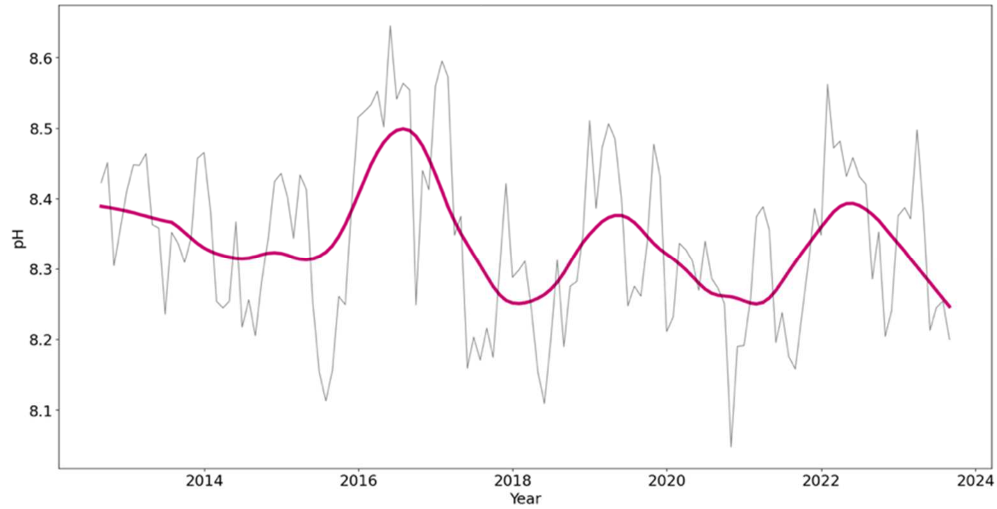

Figure 32. Monthly averaged time series (black line) and the corresponding moving average (red line) for pH.

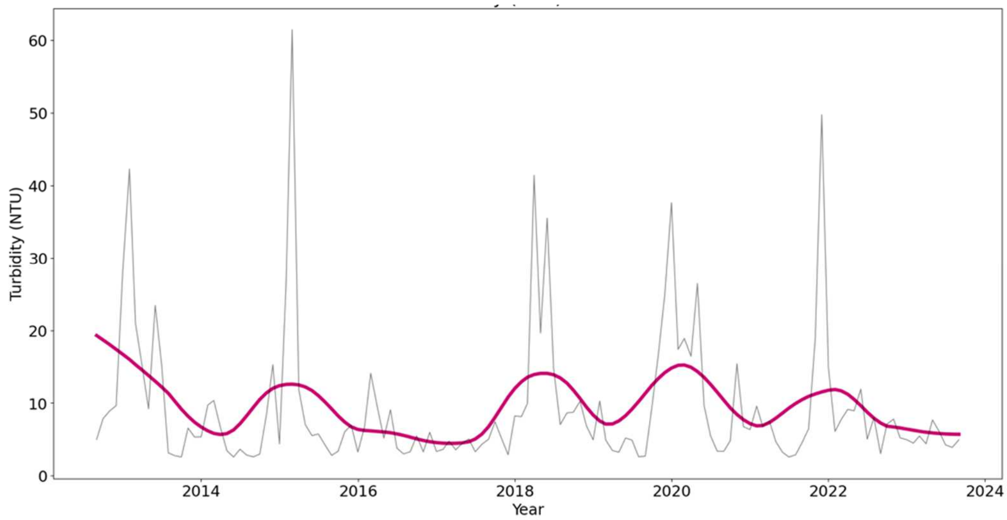

Figure 33. Monthly averaged time series (black line) and the corresponding moving average (red line) for turbidity.

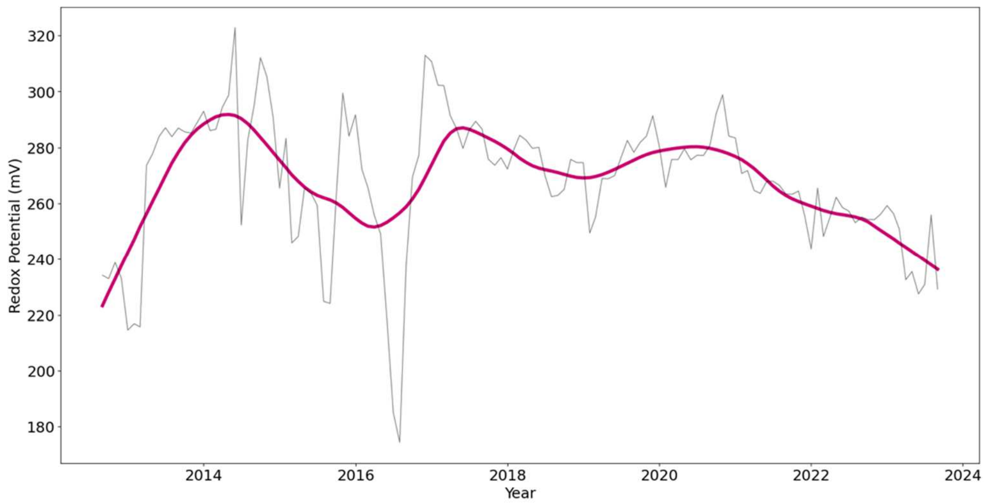

Figure 34. Monthly averaged time series (black line) and the corresponding moving average (red line) for redox potential.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

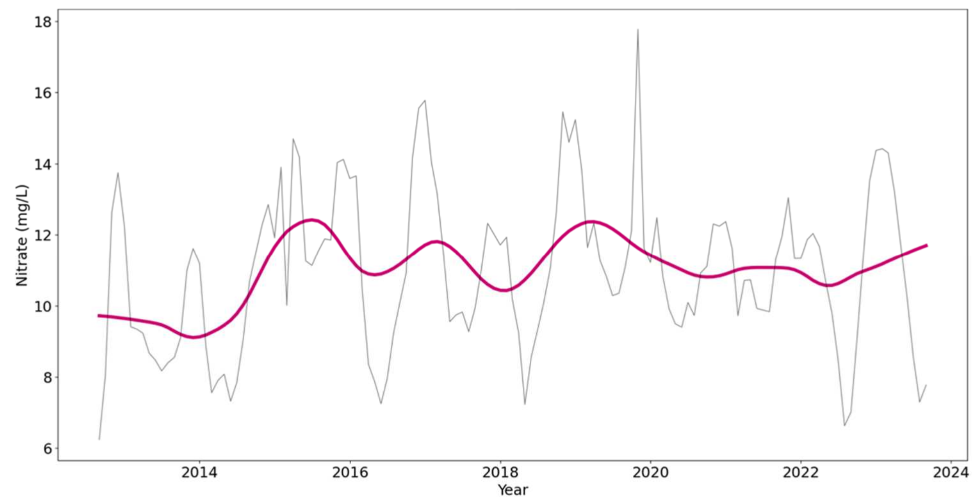

Figure 35. Monthly averaged time series (black line) and the corresponding moving average (red line) for nitrate.

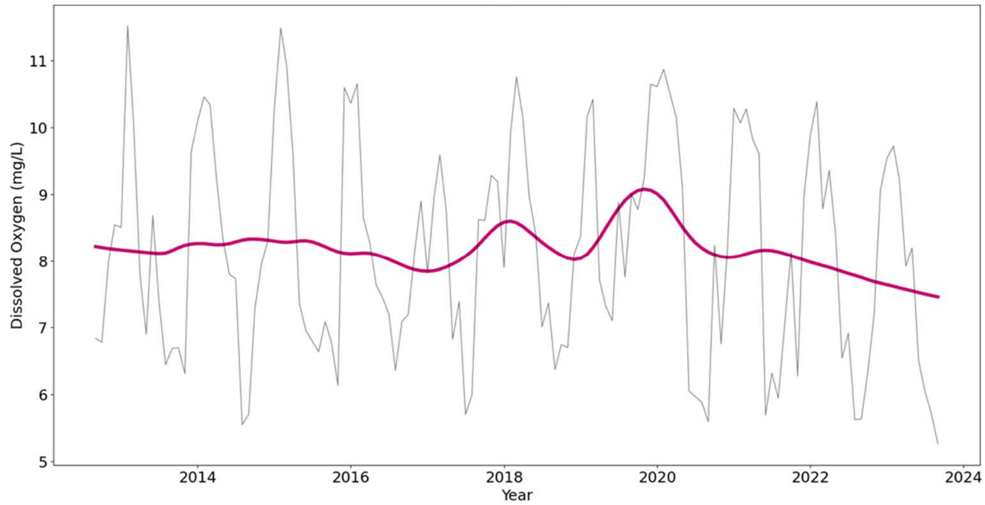

Figure 36. Monthly averaged time series (black line) and the corresponding moving average (red line) for dissolved oxygen.

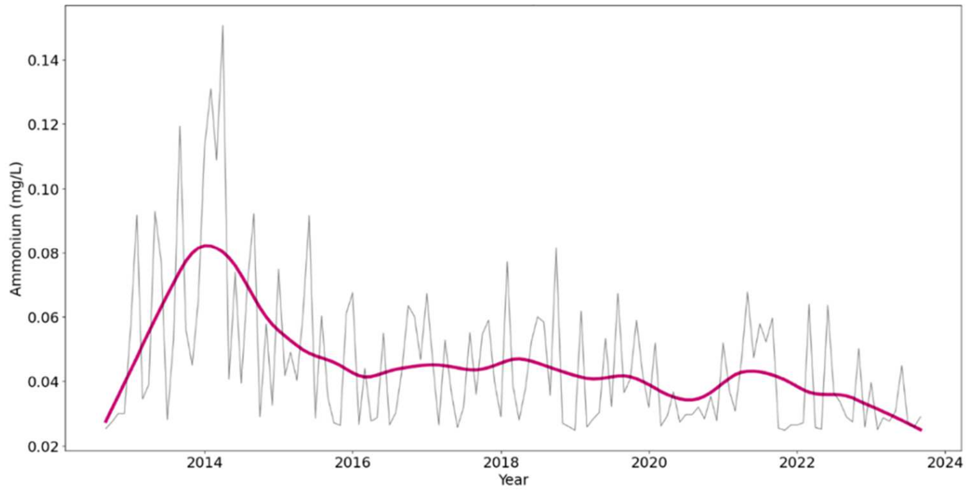

Figure 37. Monthly averaged time series (black line) and the corresponding moving average (red line) for ammonium.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

Annex 2

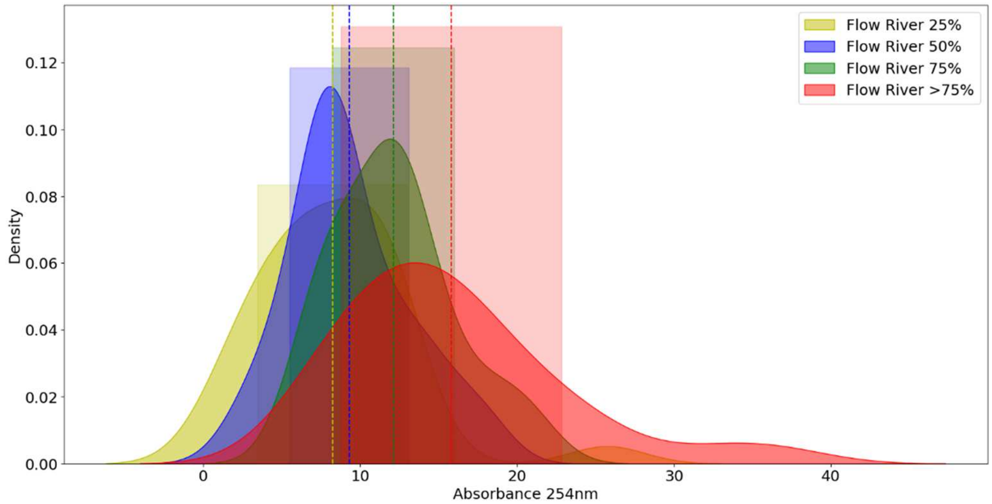

Figure 38. Absorbance at 254nm distribution divided into groups based on River flow quartiles (with mean and variance).

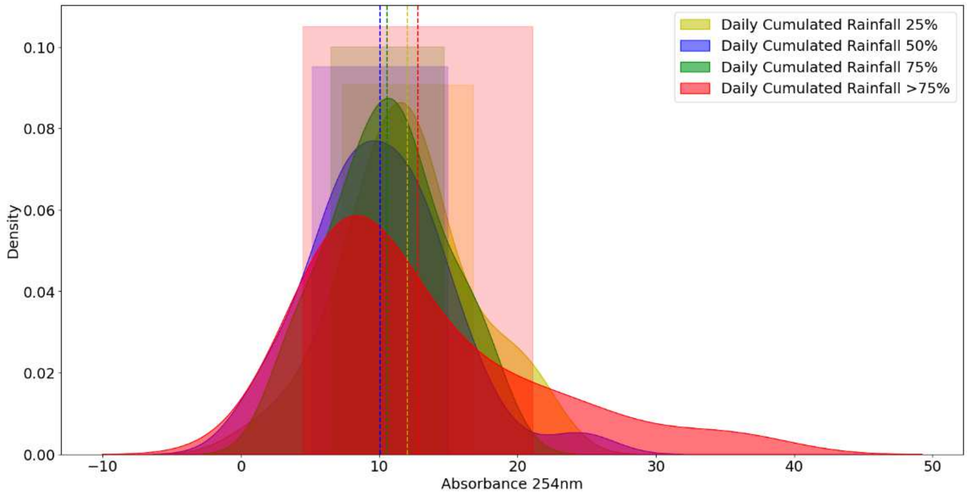

Figure 39. Absorbance at 254nm distribution divided into groups based on Rainfall quartiles (with mean and variance).

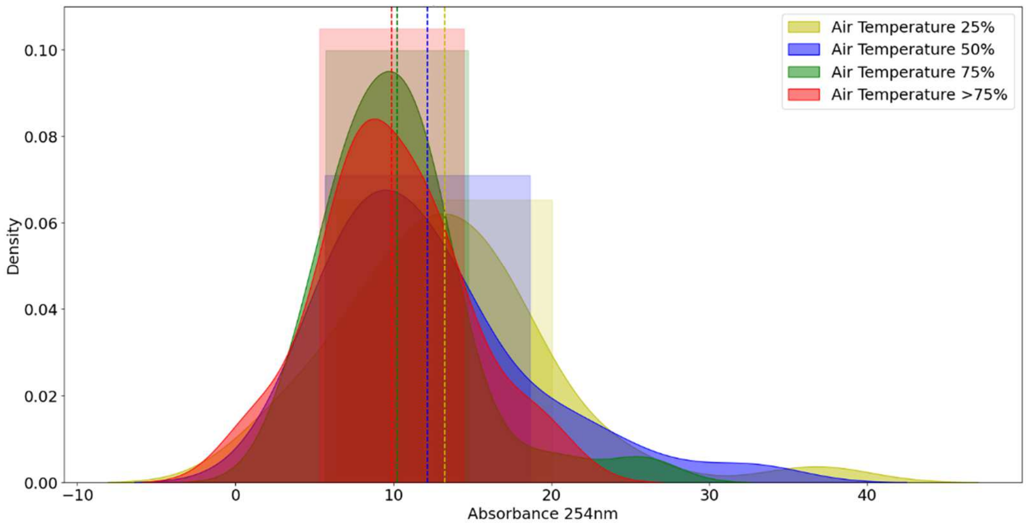

Figure 40. Absorbance at 254nm distribution divided into groups based on Air Temperature quartiles (with mean and variance).

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

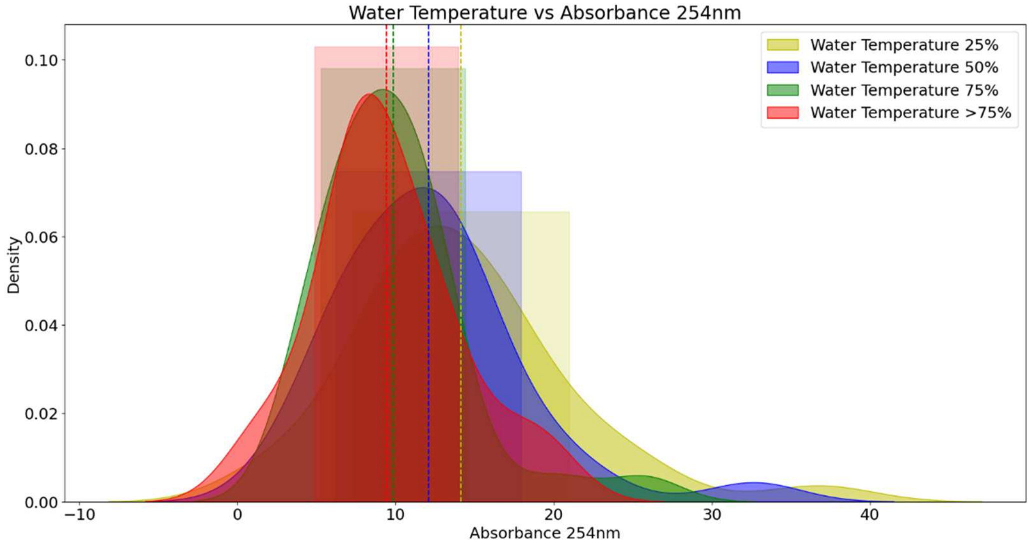

Figure 41. Absorbance 254nm distribution divided into groups based on Water Temperature quartiles (with mean and variance).

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

Annex 3

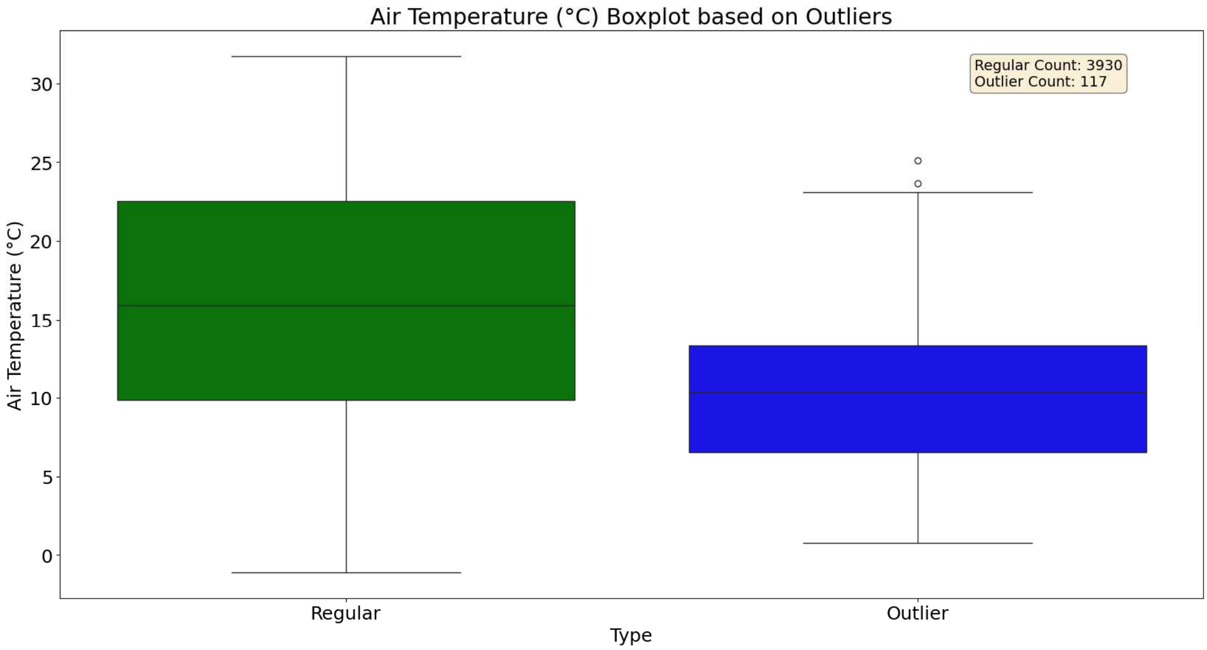

Figure 40. Air Temperature boxplots corresponding to absorbance 254nm normal and outlier values.

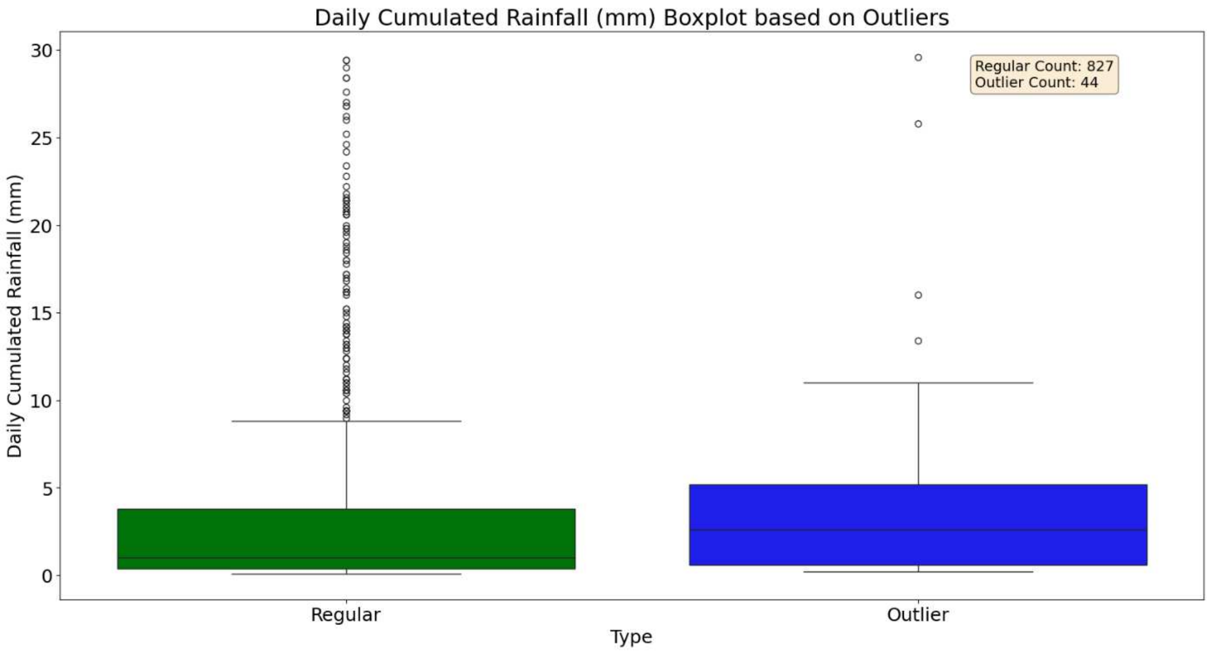

Figure 42. Rainfall boxplots corresponding to absorbance 254nm normal and outlier values (only values > 0 and < 30).

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.

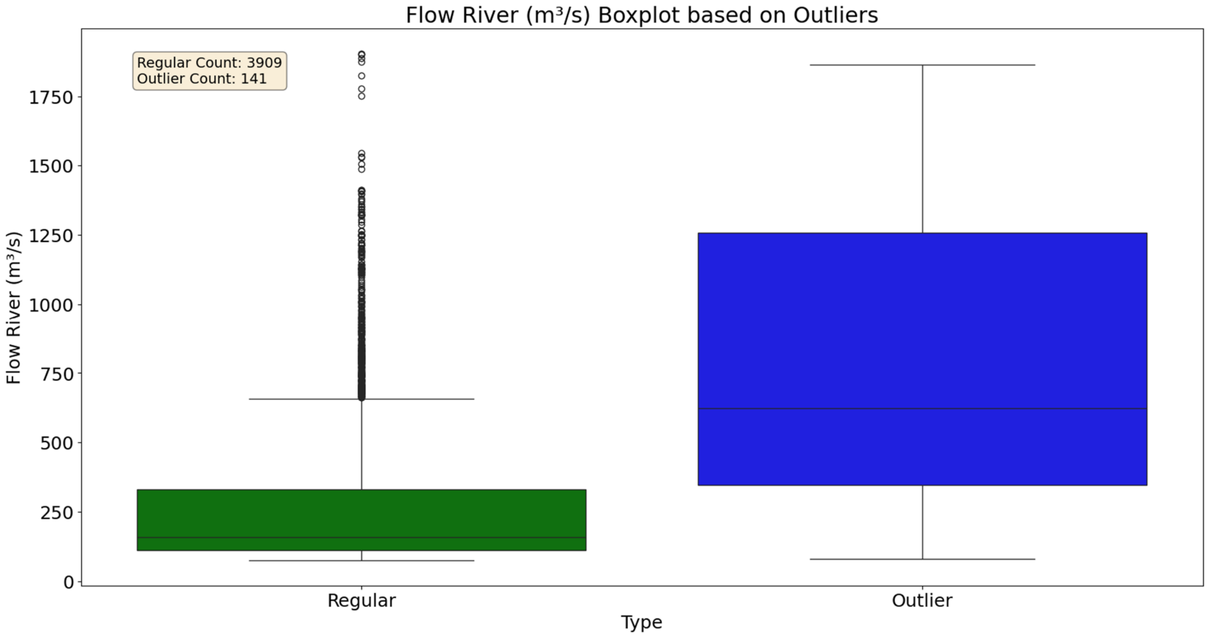

Figure 44. River flow boxplots corresponding to absorbance 254nm normal and outlier values.

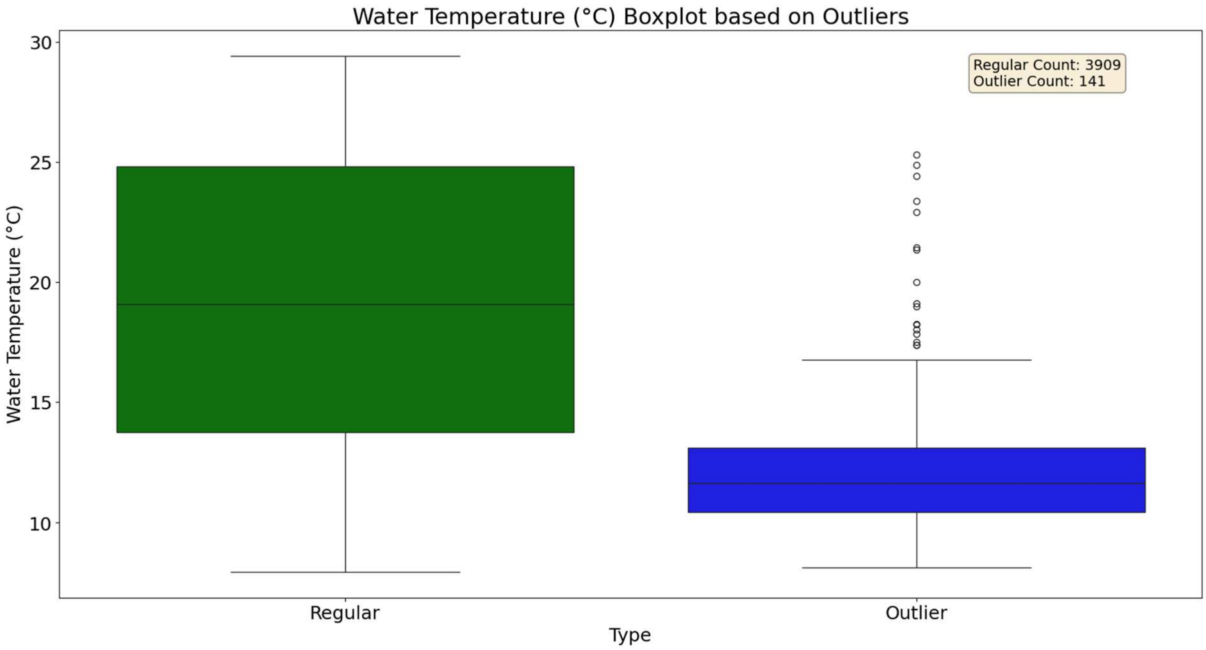

Figure 46. Water Temperature boxplots corresponding to absorbance 254nm normal and outlier values.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081980.