`

THESIS

AI/ML TOOLS FOR EARLY DECISION MAKING IN WATER SYSTEM OPERATIONS: MANAGING NON-STATIONARITY WATER QUALITY

Submitted by

Guillermo Alonso Vizarreta Luna

Department of Systems Engineering

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Summer 2025

Master's Committee:

Advisor: Steven Conrad

Mazdak Arabi
Neil Grigg
Alan Kennan

`

ABSTRACT

AI/ML TOOLS FOR EARLY DECISION MAKING IN WATER SYSTEM OPERATIONS: MANAGING NON-STATIONARITY WATER QUALITY

The assumption that natural systems oscillate within a stationary range of variability has traditionally guided water system management and allowed water utilities to experience steady state operations. These steady state operations are viewed as the 'normal state' of the water system. However, non-stationarity events such as wildfires, droughts, and floods shift water systems to new 'states' that negatively impact water quality and complicate water treatment decision-making and performance. Many water system managers do not account for these variations systemically and instead respond reactively as watersheds shift from perceived normal states. To enhance their operational resilience and develop adaptive and robust methodologies, water utilities must gain knowledge of non-stationarity states. Artificial Intelligence (AI) and its subset, Machine Learning (ML), are emerging as key tools for addressing the impacts of non-stationarity events on water system operations.

This thesis responds to this gap by investigating how AI and its subset, ML, are emerging as key tools for addressing the impacts of non-stationarity events on water system operations. This study provides a summary and observations on: (1) Understanding the boundary influences on water quality due to non-stationarity events and their implications for drinking water treatment processes. (2) Exploring how AI/ML methods inform stationarity and non-stationarity system state patterns. (3) Applying AI/ML models to develop a TOC predictive tool and assess their potential use to address non-stationarity water quality states.

`

ACKNOWLEDGMENTS

I extend my sincere gratitude to Dr. Steve Conrad for his unwavering guidance and commitment throughout the course of this research, not only as my academic advisor, but also as a human being whose empathy and support provided clarity during my experience as an international student.

I am also grateful to my colleagues and friends at the BlueGreen Decision Lab: Josh Rodriguez, Dixie Poteet, Joshua Oluwatumise, and Graeme Troxell, for generously sharing their expertise, offering continuous assistance, and contributing to a supportive and stimulating lab environment.

To my international friends and fellow Fulbright scholars, thank you for your companionship and shared sense of purpose throughout this academic journey, especially to the Comunidad del Anillo.

I deeply appreciate the Fulbright Program and Fulbright Peru for placing their trust in me and granting me the opportunity to pursue graduate studies in the United States.

Finally, I wish to thank the team at Denver Water for their ongoing support of this research and for providing a collaborative and enriching professional environment.

`

DEDICATION

To my family and friends in Peru—thank you for the endless love, patience, and for pretending to understand what I do.

And to my dog Fiby, for being the best research assistant I never asked for (but definitely needed).

`

TABLE OF CONTENTS

`

LIST OF TABLES

`

LIST OF FIGURES

`

CHAPTER 1 – INTRODUCTION

The concept of stationarity, which establishes that natural systems oscillate within a consistent range of variability (Milly et al., 2008), is an evolving framework for water management. This supposition has traditionally guided the statistical analysis of hazards in water management and design (Slater et al., 2021). However, anomalous and extreme climate events challenge the assumption that historical variability of natural systems will remain constant (Lyle et al., 2023). Climate change is manifesting noticeable non-stationarity effects by intensifying various weather patterns and resulting in more frequent or severe extreme events worldwide, including floods, heatwaves, droughts, and wildfires (Bonisławska et al., 2023; Heidari et al., 2021; IPCC, 2023; Sheffield et al., 2012; *WMO*, n.d.; Yuan et al., 2023). This results in increasingly non-stationary water quality states, making it difficult to determine whether fluctuating conditions represent normal system dynamics or indicative of a new system state.

Besides non-stationary abnormal events, water quality management usually faces indirect threats from these phenomena like eutrophication, oxygen depletion, and water stratification (Terry & Lindenschmidt, 2023). Moreover, post wildfire soil and debris erosion can overload the performance of the Drinking Water Treatment Plants (DWTPs) due to high and continues turbidity levels (Rubio-Martin et al., 2023). These normal and abnormal water quality disruptions add uncertainty and layers of complexity to drinking water systems, complicating decision-making processes for water utility operations.

In practice, water management decisions often rely on understanding the tails of the streamflow distribution. Having more comprehensive data improves our ability to interpret these tails, thereby informing decisions such as setting water quality permits for point-source discharges

(based on low-flow conditions) or establishing load allocations for nonpoint source pollution (based on high-flow events). However, as systems shift toward non-stationarity, interpreting these tails becomes increasingly challenging due to growing uncertainty about future hydrologic conditions and evolving statistical properties that can defy water managers' expectations (Hirsch, 2011).

Addressing non stationarity necessitates access to reliable and comprehensive water quality data. These data are essential for two reasons. First, it enables the development of precise indicators that underpin conventional water planning and management strategies—crucial for designing, constructing, and upgrading treatment and distribution facilities (Behmel et al., 2016; Camara et al., 2020; Richards et al., 2023). Second, robust data collection supports the integration of computational technologies such as Artificial Intelligence (AI) and Machine Learning (ML), which could aid the management of the complex and dynamic patterns of non-stationary water quality, particularly in the wake of extreme climate events.

AI is a rapidly growing area of research and application in water management. AI encompasses a wide range of computational methods, from Deep Learning (DL) to rule-based symbolic logic, that significantly improve human-machine skills and performance (Rouse, 2020). AI is frequently applied approach to incorporating predictive tools into intelligent decision support systems (Cody et al., 2020) like for urban water systems. AI has been applied in the management of water conservation projects to accelerate project implementations (T. Shi & Wu, 2021), in water quality predictions that can be used to as early warning system to prevent future water conditions and develop effective management strategies (Q. Zhang & You, 2024) as well as providing low-cost systems to monitoring and detecting abnormal patterns that may represent environmental and human health risks (Czyczula et al., 2022). Given the strength of AI approaches for analyzing large

and complex data sets, these techniques could present solutions for identifying and understanding non-stationarity water quality states.

1.1 Research Objectives

This study focuses on investigating methods for identifying non-stationarity in water systems. As climate change drives water systems toward potential shifts in non-stationary water quality states, this research aims to understand the behavior of water quality. Specifically, this research focuses on Total Organic Carbon (TOC), a key parameter chosen because of its role in the formation of disinfection by-products (DBPs) and its importance for regulatory compliance and human health.

The changing dynamics of TOC were also investigated within the North Fork of the South Platte River as it flows through Strontia Reservoir, a critical drinking water source for the city of Denver, USA. This research analyses TOC residuals to assess if they can represent potential indicators of extreme weather events that represent a new recurrent pattern of abnormal TOC in the system.

To predict TOC this research also explores the viability of AI and ML models to predict TOC trends and evaluate AI/ML abilities to understand non-stationary conditions. Whereas applications of AI tend towards optimization of predictive models, the aim of this research is to provide insights into TOC non-stationarity behavior and assess AI/ML as a predictive tool that can be applied in non-stationarity scenarios to enhance water quality decision-making in Denver Water operations.

1.2 Contributions to the body of knowledge

- Investigating how external influences affect water utilities and exploring how AI/ML-driven approaches can mitigate these challenges.

- Analyzing the temporal variability of Total Organic Carbon (TOC) in the North Fork of the South Platte River, Colorado, USA, under changing hydrological and climatic conditions.
- Assessing TOC residuals to determine if they can represent a new recurrent pattern of abnormal TOC that contributes to non-stationarity in the system.
- Exploring how AI/ML techniques can enhance Denver Water operations by assessing AI as a TOC modeling tool and assessing their capability to handle non-stationary patterns through the integration of TOC residual analysis with AI-driven models.

`

CHAPTER 2 – BACKGROUND

The performance of DWTPs is influenced by various factors, including the raw water quality, appropriate chemical dosing, the hydraulic characteristics of each treatment step, staff commitment to achieving the final goal, and the quality of the treated water (Al-Obaidi et al., 2020; Melo et al., 2016). According to Melo et al., (2016) and Zhang et al., (2012), optimal DWTP performance hinges on three key factors: robustness, reliability, and resiliency. Robustness ensures the plant consistently produces high-quality effluent despite seasonal changes in raw water quality, reliability measures the likelihood of the plant meeting water quality standards set by regulatory agencies or internal benchmarks over time, and resiliency refers to how quickly the plant recovers and resumes high-quality production after a system failure, such as an interruption in coagulant addition.

To understand the relationships and interactions between water quality and water utility operations, a literature review was conducted using Web of Science and its indexed database collections. The insights gained from this review were used to develop a conceptual model, grounded in a systems thinking map, employing causal-loop techniques. This conceptual model aims to better understand how extreme weather events impact drinking water treatment processes and provide a comprehensive overview before applying the AI/ML models.

2.1 Understanding boundary influences in water quality and their implications in water utility operations

Key water quality parameter for DWTPs would be addressed in this section and explain their importance during treatment operations.

2.1.1 Raw water quality operation parameters

Raw water quality dictates the operation of DWTPs. Coagulation-flocculation, sedimentation, filtration, and disinfection are considered the conventional phases of drinking water treatment for surface water (Rizzo et al., 2004). The operational performance of these phases is related to water quality parameters, which largely depend on the design and scale of the treatment plant. As a result, determining the maximum levels that various treatment phases can handle can be challenging, as these capacities vary from plant to plant and because fluctuations in certain parameters can synergistically affect others. Therefore, given this variability, key operational water quality parameters are outlined and analyzed to understand the impacts on different treatment phases.

2.1.1.1 Turbidity

Turbidity is an important water quality measure due to its effect on treatment processes and bacteria and virus formation. Water turbidity is caused by the accumulation of fine suspended particles and colloids from organic or mineral material (Szpak et al., 2020). Increased turbidity levels are often attributed to factors such as higher flow from spring snowmelt, heavy rainfall, and bank erosion (Gauthier et al., 2003). When raw water turbidity rises, it can adversely affect critical treatment processes, including coagulation, flocculation, sedimentation, dissolved air flotation, and filtration (Nemani et al., 2023). According to Khan et al., (2015), sand filtration and conventional coagulation are less effective in meeting optimum water quality standards when there are high levels of turbidity. Szpak et al., (2020) in collaboration with a water utility, determined that when river water turbidity reaches 1,000 NTU, a decision should be made on whether to suspend water extraction until conditions return to acceptable levels. An increase in raw water turbidity has been shown to negatively affect the disinfection process by shielding bacteria and viruses from disinfectants (Chen et al., 2017; Szpak et al., 2020). Studies have also demonstrated

a positive correlation between raw water turbidity and the presence of Giardia and Cryptosporidium (Nemani et al., 2023).

2.1.1.2 Total Organic Carbon (TOC)/ Dissolved Organic Carbon (DOC)

A major challenge in consistently supplying safe and clean water is the presence of natural organic matter (NOM) in raw water sources (Assefa et al., 2024). NOM is a complex mix of organic substances found in all natural waters and is classified as either particulate or dissolved, depending on whether it is retained by a 0.45 μm filter (Menya et al., 2018). The particulate fraction is related to TOC, while the dissolved fraction is related to DOC. TOC and DOC are widely measured parameters that provide a reliable assessment of NOM removal in drinking water treatment processes (DeMont et al., 2024; Millar et al., 2016; Shetty & Goyal, 2022; Sillanpää et al., 2015). Furthermore, elevated levels of TOC can also promote algal growth impacting water treatment costs (H. Guo et al., 2022).

Research has long identified a relationship between high levels of NOM, TOC, and DOC and the formation of DBPs. These compounds are formed when disinfectants react with NOM during the disinfection treatment phase (Assefa et al., 2024; Blackburn et al., 2023; DeMont et al., 2024; Koley et al., 2024; Samson et al., 2016; Xiao et al., 2023). Trihalomethanes (THMs) are the most produced DBPs and they are a major concern due to their carcinogenic and toxic nature, posing significant risks to human health (Albanakis et al., 2021; Guinea et al., 2024; Yirenkyi-Fianko et al., 2022). THMs such as chloroform ($CHCl_3$), bromodichloromethane ($CHBrCl_2$), dibromochloromethane ($CHClBr_2$), and bromoform ($CHBr_3$) are particularly formed during the chlorination process (Kumari & Gupta, 2022).

2.1.1.3 Alkalinity

Alkalinity is a water quality parameter that indicates the water's buffering capacity. Buffering capacity is defined by the concentration of ions that neutralize hydrogens ions acids

(Bozorg-Haddad et al., 2021). This buffer is crucial for the overall treatability of raw water, as it determines the water's capacity to chemically react (Nabors et al., 2011). Total Alkalinity (TA) is a widely used measurement in drinking water treatment, serving as a crucial tool for interpreting and controlling treatment processes (Ahmad et al., 2023). In freshwater, TA is primarily composed of carbonate alkalinity (HCO_3^- and CO_3^{2-}), with a minor contribution from hydroxide ions (OH^-) (Lehmann et al., 2023). The concentration of alkalinity in raw water plays a key role in conventional treatment processes. Higher levels of alkalinity can hinder coagulation by the formation of less compact flocs (Cao et al., 2011; Towler et al., 2009). Furthermore, alkalinity reduces the genotoxicity of chlorinated water by up to 55%, but this comes at the cost of increased THM formation during chlorination (Fang et al., 2023).

2.1.2 Stationarity and non-stationarity implications on water systems

In this section, concepts and studies related to stationarity and non-stationarity in the water sector are discussed, focusing on their implications for water quality modeling, hydrology, and treatment operations.

2.1.2.1 Stationarity

Stationarity has been a foundation of water system design for decades. Stationarity is observed when the statistical properties of a time series event, such as the mean and variability, remain constant over time (K. W. Murphy & Ellis, 2014; Yang et al., 2021; Yuan & O'Loughlin, 2024). The assumption of stationarity has guided the water sector to function within a deterministic framework, where uncertainty is minimized, ensuring reliable management of both water quality and quantity. The planning and design of water systems assume of stationarity (Wang & Yang, 2024; Yang et al., 2021). For instance, flood frequency analysis assumes stationarity by working with statistical properties of flood historical records (Yilmaz et al., 2023) allowing the design of adaptive and mitigation flood responses.

Stationarity can be categorized into two types: strong stationarity and weak stationarity. Strong stationarity means that the joint distribution of any collection of variables in a time series remains unchanged when the time indices are shifted. Strong stationarity is a rigorous condition that it is rarely applicable in real-world scenarios, especially in time series representing environmental behaviors, such as hydroclimate or water quality in water systems. In contrast, weak stationarity indicates that the time series has a constant mean over time, and its auto-covariance depends only on the time difference (or lag) between points, not their absolute positions (Tibshirani, 2023).

For the purposes of this thesis, the concept of weak stationarity was adopted, as it is more commonly used in practice (Tibshirani, 2023; Yang et al., 2021). In this context, Raw Water Quality States (RWQS) are defined as specific water quality conditions for which DWTPs know the appropriate treatment configurations to operate effectively, even if these conditions fluctuate (under weak stationarity) in response to changes in watershed and climate factors. For example, seasonality is a key factor that drives variations in RWQS.

2.1.2.2 Impacts of seasonality on raw water quality

The operation of DWTPs differs notably from other industrial operations due to variations in raw water quality and presents treatment challenges. Water utilities manage these fluctuations by tracking seasonal trends through historical records, allowing them to adjust treatment parameters according to the specific water quality changes that each season brings. Raw water quality is influenced by seasonal changes, wet weather events, and human activities within the catchment area (Bertone et al., 2016). Seasonality can cause significant fluctuations in raw water turbidity affecting drinking treatment (Y. Zhang et al., 2021)). For example, significant TOC removal efficiency (59.19% during the dry season and 66.25% during the rainy season) and THMs concentrations (136.6 μg/L % during the dry season and 108.7 μg/L during the rainy season) has

been reported in Koka Water Treatment Plant (Ethiopia) due to seasonal variations (Assefa et al., 2024). Betasso Water Treatment Plant (Colorado, USA) has also noted challenges in meeting DBP regulations due to drastic TOC seasonal variation in raw water has been reported, specifically in spring and early summer months (Samson et al., 2016). Seasonality is one factor that affect water quality in reservoir due to the strong dependence on the river hydrological regime. Górniak, (2020) analyzed TOC dynamics in Polish reservoirs finding that higher TOC concentrations were observed in the spring and summer seasons.

A study of raw water quality of a DWTP in China showed that concentrations of Total Nitrogen and Total Phosphorus (TP) are higher in the summer, while ammoniacal nitrogen are higher during winter (L. Wang et al., 2016). Krbavčić et al., (2023) observed that increased levels of manganese (Mn), iron (Fe), and ammonium (NH4) at lower pH values were found in the raw water that feed the DWTP of Butoniga during summer. Moreover, Szpak et al., (2020) found that the highest turbidity values were recorded in May, June, and July, periods that coincide with intense seasonal rainfall. Skaland et al., (2022) reported that higher maximum temperatures did not show a significant correlation with changes in any outcomes over the course of the year raw water.

A common pattern in certain areas is the contribution of snowmelt to seasonality, the increase of annual mean temperature by 1 °C from 1900 to 2014 has been related to higher runoff during winter and spring leading to premature snowmelt (Skaland et al., 2022). Samson et al., (2016) found that TOC significantly change over season when snowmelt occurs getting peak concentration during spring and early summer. Górniak, (2020) also mentioned that spring season is affected by snow melting. Water utilities manage these fluctuations by tracking seasonal trends

through historical records, allowing them to adjust treatment parameters according to the specific water quality changes that each season brings.

2.1.2.3 Non-stationary

Milly et al., (2008) started a controversial debate into the water sector by declaring that "Stationarity is dead: Whither Water Management?" They argued that Anthropogenic Climate Change (ACC) is altering the water cycle and water utilities operations due the shifting of means and extremes of precipitation, evapotranspiration and rates of discharge of rivers. Singh et al., (2024) support this idea by mention that ACC is contributing with more uncertainty and issues to the assumption of stationarity. Chen et al., (2024) notes that surface water quality is deteriorating and exhibiting increasingly non-stationary behavior due to the combined effects of human interventions and ACC. Recent studies have also agreed that the assumption of stationarity cannot be considered by default for the water sector in a changing environment anymore (S. Chen et al., 2024; Lei et al., 2021; W. Shi & Xia, 2017; Yang et al., 2021; Yuan & O'Loughlin, 2024). Figure 1: represent illustrations of time series showcasing the characteristics of (a) stationarity, along with different types of non-stationarity: (b) trend stationarity, (c) level stationarity, (d) heteroscedasticity, and (e) difference stationarity. The solid and dashed black lines denote the mean and variance functions of the time series, respectively (Salles et al., 2019).

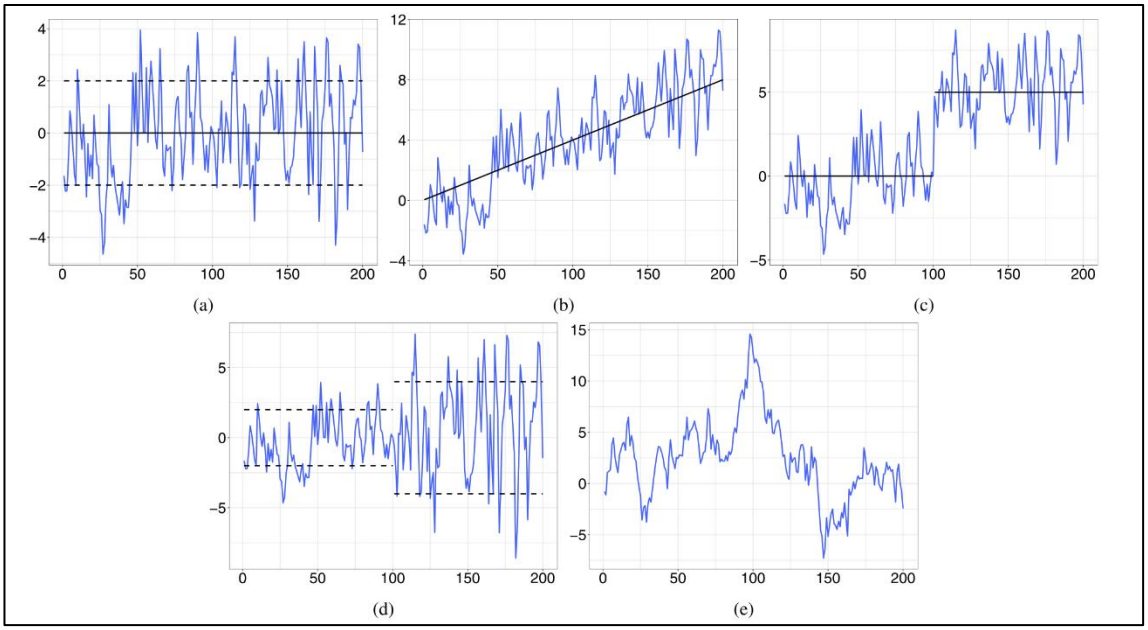

Figure 1: Examples of Stationary and Non-Stationary Time Series (Salles et al., 2019).

Despite changes in water systems, such as ACC, acidification, and land use changes, which significantly impact DWTPs (DeMont et al., 2024), the study of stationarity and non-stationarity in water quality remains limited. Shi & Xia, (2017) have examined non-stationarity, with their work identifying non-stationary patterns in ammonium nitrogen and permanganate index series in the Huai River (China) using a time-varying copula model. Additionally, DeMont et al., (2024) reported that shifts in NOM linked to climate change have been observed in surface waters across the Northern Hemisphere.

Several studies have emerged in response to the claim that "Stationarity is dead." Some support this assertion by presenting evidence of non-stationarity in hydrological patterns, while others argue that many patterns remain largely stationary. Studies have found that human interventions, such as the construction of reservoirs, are often the primary drivers of these changes. Yuan & O'Loughlin, (2024) explored non-stationarity in the Wei River and found that Baojixia diversion was the primary driver of non-stationarity in the basin. Zhang et al., (2014) used the Pettitt technique and observed non-stationarity in the Pearl River Basin caused by the presence of

abrupt change points in the time series of annual peak flood records due to the construction and operation of reservoirs. Wang & Yang, (2024) analyzed streamflow observations from 22,000 catchments worldwide and found that historical annual records has mainly stationarity in most of the catchments exclusively affected by climate change, while annual streamflow observations are mostly non-stationary in catchments mainly impacted by human interventions. Murphy & Ellis, (2014) found evidence of non-stationarity in temperature, while precipitation and runoff exhibited stationary behavior in watersheds of the Colorado River Basin. In the Poyang Lake Basin (China), stationary characteristics were predominantly observed in extreme precipitation indices, while significant non-stationary patterns were identified in the duration of extreme precipitation events (Lei et al., 2021). These studies do note that there is a lack of sufficient data to conclude whether water systems are experiencing non-stationarity, and the results are not universally applicable.

Direct relationships between climate and water quality remain largely unknown on a global scale due to the limited availability of studies specifically focused on the impact of climate change on water quality (J. Raseman et al., 2017). Closing this gap depends on statistics-based metrics for predicting future extreme events. These predictions depends on the assumption of stationarity, which presumes that future variations in extremes will follow historical patterns. However, if historical extreme events display nonstationary trends, metrics derived from this historical data may fail to effectively capture future changes in extremes (Wang & Yang, 2024). This is because such metrics do not account for new patterns or unprecedented extremes that have not been observed before. The paradigm of stationarity is shifting to non-stationarity factors driven by human interventions that are easier to predict their impact such as land-use change and water resources extraction, and driven by more complex processes such the hydrologic cycle associated to ACC (Serinaldi et al., 2018). Future research is needed to distinguish between these factors and

to analyze the frequency of such events, helping to understand how water quality is shifting to new states.

Conventional statistical trend and change point tests are effective at identifying non-stationarity in hydroclimatic indices with low variability, such as temperature. However, these tests struggle to detect non-stationarity in more variable processes, such as rainfall and streamflow (Singh et al., 2024). To address non-stationarity, we need to find methods that can identify and manage these shifts (Milly et al., 2008; Q. Zhang et al., 2014). AI, ML and DL models present areas of research. These models can help improve our understanding and predictions of water quality dynamics, supporting the effective management of water systems (S. Chen et al., 2024; Reichstein et al., 2019).

Nevertheless, we are not discarding stationary models. Nevertheless, water managers are not discarding stationary models. These models remain effective and will continue to work in areas where hydrological, climate variables, or water quality have not yet shifted and may not shift for a long time. However, where there is evidence that patterns have shifted, such as the occurrence of extreme weather events or outcomes indicating a lack of stationarity, then non-stationary models need to be incorporated and used to complement traditional stationary approaches. This combined approach will enable us to successfully address these abnormal changes in water systems.

When extrapolating literature findings to water quality in the South Platte watershed, the study area, both present and future conditions are shifting toward non-stationarity. Climate change, with its increasing frequency and intensity of extreme weather events, is emerging as a major driver of this shift, potentially equal to or even greater than the impact of human interventions. For instance, Environmental Protection Agency (EPA) study used climate models and watershed simulations to project significant hydrologic changes in the South Platte watershed, projections

showed a 10% decrease in streamflow volume, a 5% reduction in the 7-day low flow, a 29% increase in the 100-year peak flow, and a 13-day shift toward earlier season runoff (OCIA, 2015).

The focus of this thesis thus assumes non-stationarity is present in the study watershed, as it explores the connection between non-stationarity in raw water quality—caused by extreme events. However, due to the limited availability of research on this specific topic, I draw from literature that examines non-stationarity in hydrological variables, such as streamflows, where more studies are available. Literature notes connection between water quality and streamflows, both of which can be affected by human activities or intensified by climate change (Dallison et al., 2022, 2022; D. Guo et al., 2019; Tu, 2009). Therefore, the following non-stationarity events in hydrological systems can be extrapolated to water quality.

2.1.2.4 Extreme weather events

The non-stationarity nature of extreme weather events has been attributed to climate change effects. The higher frequency of extreme weather events caused by climate change can change surface water quality critically posing operational challenges to DWTP (Khan et al., 2015; Nemani et al., 2023; L. Wang et al., 2016; Xiao et al., 2023). Most DWTPs have been designed based solely on historical peaks in water quality parameters, and may not be adequately equipped to handle the intensified extremes driven by climate change (Nemani et al., 2023). Climate change and extreme weather events contribute to the non-stationarity of RWQS and are associated to water quality degradation in catchments (J. Raseman et al., 2017) impacted by elevated levels of suspended particles, organic matter, nutrients, inorganic compounds, and pathogenic microorganisms (Skaland et al., 2022).

Therefore, extreme weather events have diminished water utilities' capacity to provide high-quality potable water due to both the immediate and long-term effects. These events are observed at various stages of the treatment processes, including before and after treatment (Bertone

et al., 2016). According to (Konapala et al., 2020) shifts in precipitation patterns and increasing water demand due to climate change will affect global water availability, this will make even more challenging to provide drinking water for a growing population. In a review of 46 case studies from Australia and the USA, heavy rainfall was the most reported extreme weather event, with the broadest range of potential impacts on water utilities (Khan et al., 2015; Wright et al., 2014).

Extreme precipitation events are also expected to increase sedimentation and contamination runoff (Bertone et al., 2016; Xiao et al., 2023; Younos & Grady, 2014). This surface runoff is associated with higher levels of natural organic matter (NOM and DOM) and increased transport of TOC to water sources, necessitating the implementation of costly and complex advanced treatment methods (Samson et al., 2016). Additionally, this situation requires an increase in chlorine disinfection doses, which in turn leads to higher formation of DBPs (J. Raseman et al., 2017; Khan et al., 2015). Increasing of TOC also puts pressure on coagulation and sedimentation processes, potentially compromising their ability to maintain optimal removal efficiency(Xiao et al., 2023). Increased runoff due to extreme precipitation are associated with increased levels of Escherichia coli (E. coli), coliform bacteria, intestinal enterococci, color and turbidity in raw water (Skaland et al., 2022). For example, after the occurrence of extreme precipitation events such as like typhoons, higher values of microbiology parameters were founded such as 5000 MPN/100mL of total coliforms and 880 MPN/mL thermotolerant coliforms (W. Chen et al., 2017). While during seasonal monsoon events, elevated turbidity and higher microbiology parameters were also founded impacting drastically water quality (Koley et al., 2024).

2.1.2.5 Other events

Wildfires are often present in drinking water watersheds are one of the current critical environmental issues worldwide because they degrade water quality supply sources and air quality both impacting public health. There is a connection between the occurrence of wildfires that

release particles, such ash, into the soil, and the subsequent precipitatTion that transport these particles to the water sources degrading water quality (Blackburn et al., 2023). Key factors that will define the degree of impacts in water quality are the duration period of a fire and the timing and magnitude of the next precipitations events (Bertone et al., 2016). For example, after the major wildfires in Colorado, water turbidity and nitrate levels were elevated for at almost five years (Khan et al., 2015).

Another extreme weather event that is becoming more frequent is heat waves. These events cause significant temperature increases, which can trigger algal blooms and cyanobacteria outbreaks, leading to higher levels of organic matter in water. This primarily affects the filtration treatment stage by requiring more frequent backwashing and reduces the effectiveness of disinfection processes, resulting in the creation of more DBPs (Xiao et al., 2023). Additionally, changes in temperature negatively impact TOC removal efficiency (Assefa et al., 2024) that are also related to DBP formation.

Moreover, snowmelt events, while short-term, may have long-term impacts on water quality due to climate change. Snowpack melting is expected to accelerate, and many water supply companies that depend on glaciers or seasonal snowfall may experience significant changes in water quality. For example, snowmelt is the main driver of TOC during spring and early summer in Colorado (Samson et al., 2016).

2.1.3 Mapping non-stationarity impacts on drinking treatment processes and potential benefits of AI implementation

To explore the non-stationarity impact on drinking treatment processes and potential benefits of AI implementation a conceptual model, represented by a systems thinking map, was developed. Figure 2 illustrates the interconnected impacts of climate change, extreme weather events, and water quality dynamics on drinking water treatment. It emphasizes key raw water

quality parameters such as TOC, turbidity, and alkalinity, which are essential for effective treatment processes. The integration of an Effective AI System is shown to potentially reduce operational uncertainties through anomaly detection and predictive capabilities, enhancing decision-making and system resilience.

Figure 2 enhances understanding of system boundaries, dynamic interactions, and the cascading impacts within the system. These system boundaries focus on interconnected processes driven by anthropogenic activities (Green Boundary) and climate change (Skyblue Boundary). Both factors play a critical role in leading to extreme weather events such as wildfires, impacting watersheds and water quality sources (Yellow Boundary), and creating operational challenges for DWTPs (Orange Boundary).

The systems thinking map also illustrates two lines of causal effects: one with AI implementation and one without. In the normal state of the system (without AI implementation, represented by black arrow flows), where extreme weather events continue to intensify their impacts on DWTPs, the map reveals an increase in uncertainty within operator decision-making processes. As a result, the production of safe, treated drinking water diminishes. This exacerbates unmet demand for drinking water, particularly in developing countries, where gaps in reliable water services are already severe due to inadequate infrastructure and limited budgets for operation and maintenance.

The unpredictable behavior of wildfires and the complex interplay of vegetation, landscape, and hydrology create significant uncertainty about their impacts on water quality and water utilities operations (Hohner et al., 2019), leading to increased suspended solids, such as ash from post-fire storms, that pose challenges for water treatment by affecting turbidity standards and

raising filtration and coagulation costs (Paul et al., 2022). These conditions are shown in the interaction of Skyblue, Yellow and Orange systems boundaries.

Besides the impacts on water quality and DWTPs operations, there are many different impacts associate with non-stationarity events. After the occurrence of extreme weather events gastrointestinal illnesses disruptions has been reported (Khan et al., 2015). Wastewater overflows from combined collapsed sewer systems caused by abnormal rainfall events resulted in flooding (EPA, 2004; Mott MacDonald et al., 2024; Younos & Grady, 2014).

So, as water quality degrades, water utilities consume more energy for treatment, leading to a long-term feedback loop: increased energy use to treat poorer water quality results in higher greenhouse gas emissions, which contribute to climate change. This, in turn, drives more extreme weather events that further degrade water quality. The fluctuations in surface water quality have led to increased costs and altered treatment protocols for making water safe for consumer consumption, particularly concerning parameters such as pH, alkalinity, and temperature (Mensah-Akutteh et al., 2022), even a 1% reduction in raw water quality turbidity can lead to a 0.3% decrease in chemical treatment costs, resulting in annual savings of $121 to $13,060 (2015 USD) for DWTPs (Price & Heberling, 2018).

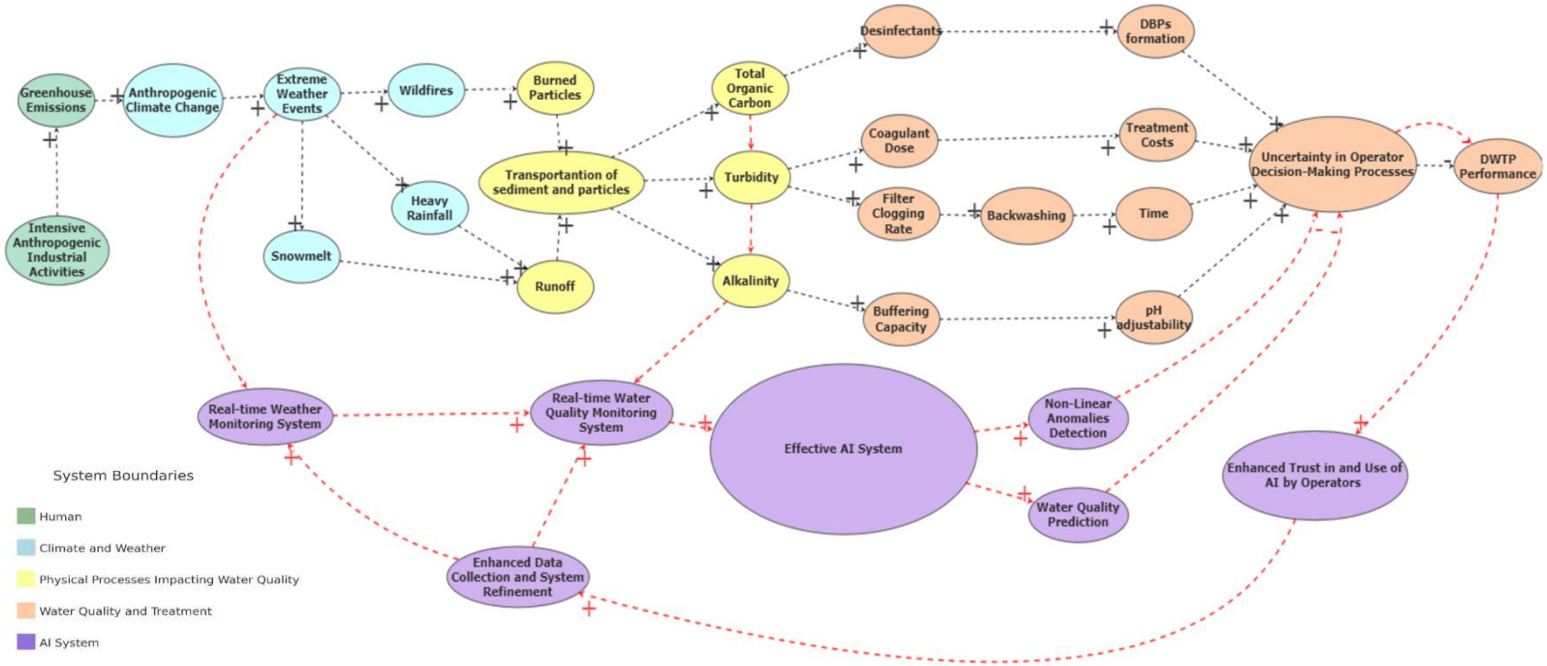

Figure 2: Systems Thinking Map.

Consequently, utilities may increase customer charges, creating inequitable access to safe water, particularly for vulnerable populations. Furthermore, polluted water sources and the limitations of DWTPs—which are not designed to handle extreme and abnormal water quality values—may result in treatment shutdowns, further aggravating water scarcity. These interdependencies highlight the urgent need for adaptive and dynamic systems to address these challenges effectively.

On the other hand, in a more resilient state of the system (with AI implementation, represented by red arrow flows), the integration of the AI boundary enhances system performance by providing real-time monitoring and predictive capabilities, thereby reducing uncertainty in operator decision-making processes. AI's ability to identify non-linear anomalies and predict water quality under non-stationary conditions ensures improved DWTP performance.

While AI can predict water quality reasonably well under "steady" conditions, the complexity and non-stationarity introduced by extreme weather events require real-time tracking. Real-time data collection allows for the capture of these rare and abnormal observations, enabling AI models to learn and adapt to the unique challenges posed by such events. Without this critical input, the models may struggle to accurately predict water quality during extreme conditions, limiting their overall effectiveness in dynamic scenarios. Therefore, AI models cannot effectively interpret extreme or abnormal changes in water quality if the training data lacks sufficient examples of abnormal patterns for the models to learn from. This highlights the importance of feeding AI systems with real-time data to enhance their performance and learning capacity continuously.

CHAPTER 3 – METHODOLOGY

This chapter details the process of collecting water quality and weather data, the methodologies used to assess the stationarity of TOC and applying feature selection techniques, as well as the foundational principles and data structuring for AI/ML models in TOC prediction.

3.1 Study Area

The South Platte River Watershed serves as the study area for this thesis. The South Platte River Watershed spans approximately 62,940 km² across three states, with the majority located in Colorado (79%), followed by Nebraska (15%) and Wyoming (6%) (Kimbrough & Litke, 1998). Originating from the Continental Divide in the central Colorado Rockies, the river flows northeast for approximately 725 km across the Great Plains before merging with the North Platte River in North Platte, Nebraska (J. Murphy & Sprague, 2019). As a critical resource, the South Platte River supports drinking water supply, recreation, and economic development, benefiting both upstream communities and the Denver metropolitan area (Eriksen & Coelho, 2017).

This thesis focuses on the North Fork of the South Platte River (NFSPR), a vital tributary within the Upper South Platte River Watershed. The NFSPR plays a key role in sustaining Denver Water's infrastructure and is a high-priority area for water quality assessment. The Upper South Platte River Watershed is a major source of drinking water, with approximately 80% of the water consumed by 1.4 million residents in the Denver metropolitan area flowing through this system (Jones et al., 2017).

According to (Denver Water, 2015), the Upper South Platte River is primarily sustained by snowmelt and water diversions through the Roberts Tunnel. The source water area varies in elevation from 5,512 feet near Kassler, Colorado, to 14,271 feet at Mt. Evans. The regional climate

ranges from alpine forests and high-elevation grasslands to lower-elevation forests, with annual precipitation between less than 12 inches and more than 30 inches. The upper watershed spans over 2,600 square miles, extending from the Continental Divide to Strontia Springs Reservoir, southwest of Denver, with elevations ranging from 6,000 to over 14,000 feet. This area contains five major municipal reservoirs and several smaller ones, making it the largest of Denver Water's source water watersheds, supplying drinking water to nearly three-quarters of Colorado's population.

The study area includes key water quality sampling points and weather stations strategically positioned along the primary water flow path. Water is diverted from Dillon Reservoir through the Roberts Tunnel into the NFSPR, which then merges with the Upper South Platte River before reaching Strontia Springs Reservoir. From there, it is transported via Conduit 26 to the Foothills Drinking Water Treatment Plant (Figure 3), strategically positioned along the primary water flow path. Water is diverted from Dillon Reservoir through the Roberts Tunnel into the North Fork of the South Platte River, which then merges with the Upper South Platte River before reaching Strontia Springs Reservoir. From there, it is transported via Conduit 26 to the Foothills DWTPs.

Table 1. Water quality sampling points associated with their closest weather stations

Water quality sampling points	Associated weather stations
WS-NF-004, WS-BL-005	RT-WS
WS-NF-005	PLAGRACO
WS-UP-051	PLABAICO, BAILEY
WS-UP-006, WS-UP-008, WS-UP-011	PLASPLCO
WS-LP-001	ST, PLASTRCO

Note: Water quality sampling points associated with the same weather stations share identical weather data due to data availability constraints. For further details, see Data Collection and Pre-Processing Section.

This thesis focuses on analyzing the stationarity of TOC and predicting its behavior using sampling points along this water flow path. Special attention is given to the sampling point "WS-LP-001", which serves as a proxy for the treatment plant influent but under natural watershed conditions, as it is located before the water enters Conduit 26. "WS-LP-001" is situated at the outlet of Strontia Springs Reservoir and, according to Denver Water, provides the most accurate TOC representation compared to "WS-UP-008", which is located at the reservoir inlet. This suggests that the reservoir's influence on water quality is minimal, making "WS-LP-001" the optimal monitoring location for TOC analysis.

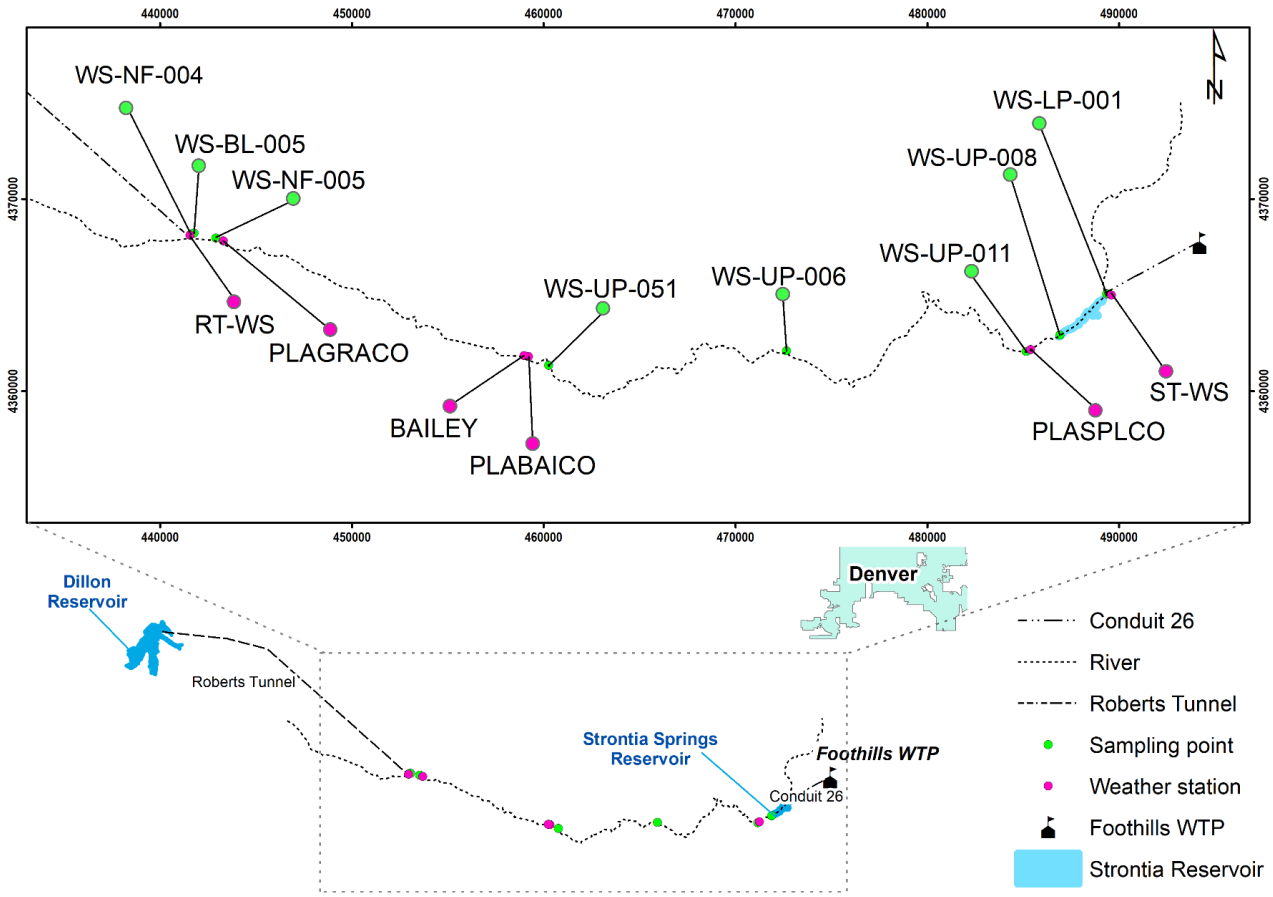

Figure 3: Study Area

3.2 Data Collection and Pre-Processing

Water quality data, collected from the Denver Water database, included over 57 water quality parameters, instantaneous streamflow (Flow_Inst) and water temperature (Water_Temp_Inst). It spans from January 2009 to December 2024, comprising 65,990 monthly observations. A 20% missing value threshold was applied; parameters exceeding this threshold were removed, while those below were imputed using historical monthly averages. Any remaining rows with missing values were eliminated. Additionally, zeros in the raw data represent values below the laboratory detection limit rather than true zero measurements. To preserve variance and avoid potential biases in modeling, parameters with a high number of zero values were removed based on expert knowledge.

Daily weather data was collected from Denver Water, Colorado Division of Water Resources, and National Oceanic and Atmospheric Administration (NOAA) stations within the study area. To ensure data quality, rows with missing values were removed. Each weather data point was linked to the nearest sampling location. Water quality and weather data were then merged based on the sampling point and sampling date, forming the Preprocessed Dataset used for the stationarity analysis and AI/ML modeling sections. After preprocessing, the dataset was reduced to 43,992 observations, containing the following measured variables:

- Metals: Dissolved (D) and total (T) concentrations of Aluminum (Al-D, Al-T), Copper (Cu-D, Cu-T), Iron (Fe-T), Manganese (Mn-D, Mn-T), Molybdenum (Mo-D, Mo-T), Nickel (Ni-D, Ni-T), Uranium (U-D, U-T), and Zinc (Zn-D, Zn-T)
- Nutrients & Organics: Nitrate + Nitrite (NO3_NO2), Total Nitrogen (TN), Total Organic Carbon (TOC), Total Phosphorus (TP)

- Ions & Conductivity: Alkalinity (Alk), Chloride-to-Sulfate Mass Ratio (CSMR), Calcium (Ca-D, Ca-T), Chloride (Cl), Conductivity (Cond), Hardness, Magnesium (Mg-D, Mg-T), Sulfate (SO4)
- Microbial Contaminants: Escherichia coli (E. coli)
- Physical Properties: Dissolved Oxygen (DO), Instantaneous Flow (Flow_inst), Total Suspended Solids (TSS), Turbidity (Turb), Water Temperature (Water_temp_inst), pH
- Weather Data: Precipitation (mm), Air Temperature (°C)

Table 2 Water quality sampling point and weather station description

Water Quality Sampling Point or Weather Station	Type of Data	Description	Data Source
WS-NF-004	Water quality, streamflow, and water temperature	North Fork of South Platte River upstream of Roberts Tunnel	Denver Water
WS-BL-005		East Portal Roberts Tunnel Outfall	
WS-NF-005		North Fork of South Platte River upstream of Geneva Creek	
WS-UP-051		North Fork of South Platte River below Bailey	
WS-UP-006		South Platte River above confluence w/ North Fork of South Platte River	
WS-UP-011		North Fork of South Platte River above confluence with South Platte River	
WS-UP-008		South Platte River at Strontia Spring Reservoir Inlet	
WS-LP-001		South Platte River at Strontia Springs Reservoir Outlet	
RT-WS	Precipitation	Roberts Tunnel Weather Station	
ST-WS		Strontia Weather Station	
PLAGRACO		North Fork South Platte River at Grant	Colorado Division of Water Resources
PLABAICO		North Fork South Platte River at Bailey	
PLASPLCO		South Platte River at South Platte	
BAILEY		North Fork South Platte River at Bailey	NOAA National Centers for Environmental Information

3.3 Analyzing stationarity and non-stationarity of Total Organic Carbon in the North Fork of the South Platte River, Colorado, USA.

3.3.1 Seasonal-Trend Decomposition using LOESS (STL)

From the Preprocessed Dataset, STL was applied as an initial step to visually assess non-stationarity in the TOC time series across all sampling points. This technique allows to clearly identify and understand underlying trends, seasonal and extreme or abnormal patterns (Equation 1). STL is filter method to decompound a time series in trend, seasonal and residual components (Cleveland, 1979). The method employs locally weighted regression (LOESS), a non-parametric technique that fits simple models to localized subsets of the data, thereby smoothing the series (Cleveland & Devlin, 1988). The analysis was implemented using the *statsmodels* library in Python (Seabold & Perktold, 2010).

Equation 1: STL equation

$$Y = T + S + R$$

Source: Adapted from (Cleveland et al., 1990)

where:

- Y = Original time series at time t
- T = Trend component (captures long-term progression)
- S = Seasonal component (captures periodic patterns)
- R = Residual component (random noise or unexplained variation)

3.3.2 Statistical Stationarity Tests

The Augmented Dickey-Fuller (ADF) and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test were implemented using the *statsmodels* library in Python (Seabold & Perktold, 2010),

with the constant + trend (ct) specification to account potential mean shift and an underlying trend in the time series data.

3.3.2.1 *ADF*

The ADF test was used to statistically determine the presence of unit roots in the TOC time series across all sampling points (Dickey & Fuller, 1981). The null hypothesis (H_0) states that the series has a unit root and is non-stationary, while the alternative hypothesis (H_1) states that the series is stationary (Equation 2).

This method extends the original Dickey-Fuller (DF) test (Dickey & Fuller, 1979) by introducing lagged differences to account for higher-order autocorrelation in time series data.

Equation 2: ADF Hypotheses

$$H_0 : p = 1 \text{ (unit root, non-stationary)}$$

$$H_1 : p1 < 1 \ (stationary)$$

Source: Adapted from Dickey & Fuller, (1981)

3.3.2.2 *KPSS*

The KPSS test was applied to statistically evaluate the presence of stationarity in the TOC time series across all sampling points (Kwiatkowski et al., 1992). The null hypothesis (H_0) states that the series is stationary around a deterministic trend, while the alternative hypothesis (H_1) states that the series has a unit root and is non-stationary. The test was implemented using the *statsmodels* library in Python (Seabold & Perktold, 2010), with the constant + trend (ct).

3.3.3 Residuals Analysis

Residuals of the TOC time series were extracted from the STL, and stationarity were analyzed using ADF and KPSS again in order to determine the presence of extreme or abnormal patterns.

3.4 Developing AI/ML models to predict TOC, enhancing decision-making in Denver Water operations.

To explore the potential application of AI/ML in water utility operations, Random Forest (RF) and Artificial Neural Network (ANN) models were developed to predict TOC across all sampling points. Additionally, predictions were made for the downstream point WS-LP-001, which represents raw water entering the DWTP under watershed conditions before going into Conduit 26 as influent.

3.4.1 Feature selection

Different feature selection methods were applied to the Preprocessed Dataset, including Correlation Analysis and Exploratory Factor Analysis (EFA) to identify linear relationships and underlying patterns for selecting the most relevant inputs for predicting the target variable "TOC" ac. Additionally, Random Forest (RF) was used to account for non-linear relationships and rank feature importance. The results from each feature selection method were combined to create multiple feature subsets (

Table 3), containing the most relevant predictors for modeling TOC using RF and ANN. Additionally, if any of these feature subsets did not include the key hydrological variables Streamflow or Precipitation, new feature subsets were generated by incorporating the missing variables. This adjustment was made to evaluate the impact of these critical hydrological parameters on model performance and assess whether their inclusion improves predictive accuracy. Feature selection techniques were conducted in Python using libraries such as pandas, seaborn, and sklearn.

Table 3: Feature Subsets

Feature Subset	Description	Parameters
Baseline (All Features)	Includes all available features in the dataset.	All available features
Correlation-Based Selection	Features selected based on strong correlations with TOC.	SO4, CSMR, Flow_inst, TSS, Zn-D, Water_temp_inst, Air_temperature, Cl, Ni-D, Fe-T, Mo-D, pH, Mo-T
Factor Analysis - Factor 1 (Hardness Components)	Represents water hardness-related components.	Hardness, Ca-D, Cond, Alk, Mg-D, Mg-T, U-D
Factor Analysis - Factor 2 (Salinity and Metals)	Captures salinity and metal-related factors.	CSMR, Cl, Mo-D, Mo-T, Mg-D, Mg-T
Factor Analysis - Factor 3 (Nickel and Copper)	Includes features related to nickel and copper concentrations.	Ni-D, Cu-D, Al-D, Ni-T, pH, Al-T, Cu-T, SO4
Factor Analysis - Factor 4 (Suspended Solids)	Focuses on suspended solids in the water.	TSS, Fe-T
Top 10 Features (Random Forest Importance)	The ten most important features identified using Random Forest feature importance.	SO4, Turb, CSMR, Zn-D, Mn-D, Flow_inst, TSS, TP, Water_temp_inst, Air_temperature
Top 20 Features (Random Forest Importance)	The twenty most important features identified using Random Forest feature importance.	SO4, Turb, CSMR, Zn-D, Mn-D, Flow_inst, TSS, TP, Water_temp_inst, Air_temperature, F, DO, Fe-T, U-T, Precipitation, Cl, TN, Zn-T, Ni-D, NO3_NO2
Overlapping Features (Across Methods)	Features that appear consistently across different selection methods.	SO4, CSMR, TSS, Zn-D, pH, Water_temp_inst
Expanded Set 1 (Including Turbidity and Temperature)	Adds turbidity and temperature variables to the overlapping feature set.	SO4, CSMR, TSS, Zn-D, pH, Water_temp_inst, Turb, Air_temperature, Mn-D
Expanded Set 2 (Including Conductivity and Hardness)	Further expands the feature set by incorporating conductivity and hardness.	SO4, CSMR, TSS, Zn-D, pH, Water_temp_inst, Turb, Air_temperature, Mn-D, Hardness, Ca-D, Cond, Alk
Comprehensive Set (Top Features Consolidated)	A consolidated set of the most relevant features across different methods.	SO4, CSMR, Flow_inst, TSS, Zn-D, Water_temp_inst, Turb, Mn-D, Air_temperature, pH, Cl, Mo-D, Fe-T, Hardness, Cond, Alk

3.4.2 Modeling

3.4.2.1 Random Forest (RF)

RF is an ensemble machine learning method used for regression and classification problems (Alnahit et al., 2022; Lap et al., 2023; Yee Wong et al., 2022), based on the combination of multiple decision trees, each built using a randomly selected set of input values that are drawn independently but follow the same distribution across all trees in the forest (Breiman, 2001). RF is trained on labeled data and make classification or predictions based on the patterns learned from these labels (Alomani et al., 2022).

It utilizes bootstrap sampling, which randomly splits the dataset into homogeneous subsets that form the root nodes of individual decision trees. A randomly selected subset of samples is then used to train each tree, while the remaining samples are set aside for validation and model accuracy estimation (F. Wang et al., 2021; Xu et al., 2021). RF can effectively handle non-linear and complex relationships (Szomolányi & Clement, 2023; Zavareh et al., 2024), which are common in water quality dynamics due to the intricate interactions between watershed systems, climate, and weather factors.

Figure 4. **RF Model Architecture** diagrams the RF modeling process. For each selected feature subset from the feature selection process, the dataset was partitioned into training (70%) and testing (30%) subsets to ensure that the model was trained on a diverse set of observations while maintaining an independent test set for performance evaluation. Multiple RF models were developed, each trained on a different selected feature subset to predict the target variable “TOC”. By systematically training models with different selected feature subsets, this research aimed to

identify the most effective combination of linear and non-linear predictor variables that influence model accuracy and stability. Each model was configured with 100 trees (n_estimators=100) and a fixed random state of 42 (random_state=42) to ensure reproducibility and consistency across experiments.

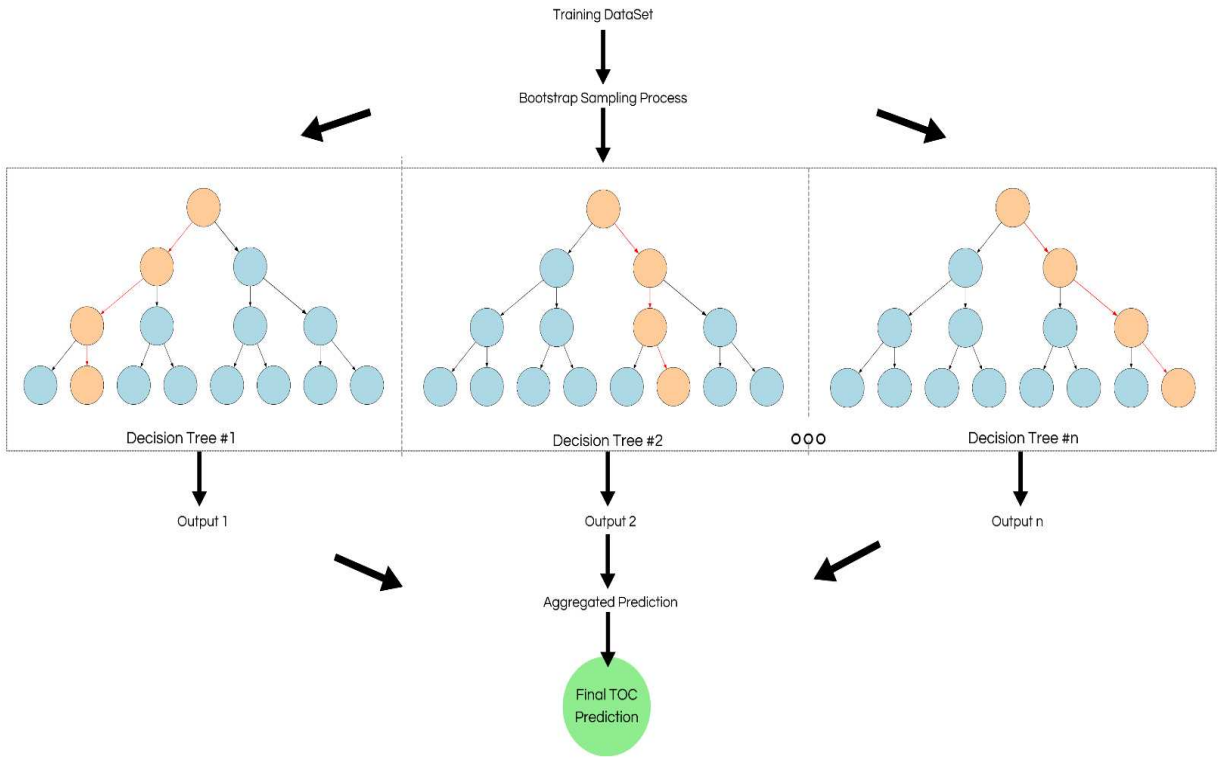

Figure 4. RF Model Architecture

Once trained, each model was evaluated based on its predictions on the test set. Performance was assessed using key regression metrics, including Mean Squared Error (MSE), Mean Absolute Error (MAE), and the R^2 score. To further evaluate the generalization capacity of each model, a 5-fold cross-validation R^2 score was computed, providing insights into how well the model would perform on unseen data. These evaluation metrics facilitated a robust comparative analysis of different feature selection strategies and their impact on predictive performance.

To identify the most effective selected feature subset, performance metrics were collected and ranked based on cross-validated R^2 scores. The two best-performing selected feature subsets were chosen for further hyperparameter tuning to enhance model performance. The hyperparameter tuning was conducted to optimize the RF models for the top two selected feature subsets, aiming to enhance predictive accuracy by identifying the best hyperparameter combination for each model.

Using Grid Search with 5-fold Cross-Validation (GridSearchCV) from scikit-learn, a predefined hyperparameter grid was tested, including variations in n_estimators (100, 200, 500), max_depth (10, 20, 30, None), min_samples_split (2, 5, 10), min_samples_leaf (1, 2, 4), max_features (sqrt, log2, None), and bootstrap (True).

For each feature subset, the dataset was split into training (70%) and testing (30%) sets, and the Random Forest Regressor was trained using GridSearchCV to identify the best configuration based on the R^2 score. The optimal models were then evaluated on the test set using MSE, MAE, and R^2 score to assess performances again. Finally, the models trained on the two best-performing selected feature subsets were retrained using the optimized hyperparameter settings and re-evaluated on the test set to confirm performance improvements.

3.4.2.2 Artificial Neural Network

ANNs are one of the earliest forms of predictive artificial intelligence techniques. ANN have been used in water resources to process both linear and highly complex non-linear patterns, including climate, hydrological factors, and water quality (Farzana et al., 2024; Haider et al., 2024; Wu et al., 2014). They have been widely applied in water quality predictions due to their capability to uncover hidden patterns and relationships in historical data (Ubah et al., 2021). Studies on apply ANNs to water quality remain within the academic sphere due to the scarcity of comprehensive

water resource data, particularly in water quality monitoring, which limits their practical application.

The foundational framework for ANN, which analyzes neural network behavior through propositional logic, was established in 1943 by McCulloch & Pitts. They conceptualized neurons as binary units that operate under an all-or-none principle, where synaptic connections correspond to logical operations such as AND, OR, and NOT (McCulloch & Pitts, 1943).

Inspired by the structure of the human brain (Chhipi-Shrestha et al., 2023; H. Guo et al., 2022), ANN architecture consists of multiple input signals—such as water quality parameters—that are received by artificial neurons. Each input is assigned a corresponding weight, which determines its impact on the neuron's response. These weights are adaptable coefficients that regulate the strength of the input signals. The neuron's output, or prediction—such as TOC in this study—is computed using a summation function that aggregates all weighted inputs, mimicking how biological neurons integrate incoming signals (Lek et al., 2008) .

The ANN modeling approach followed the same methodology as the RF models, with the primary difference being the use of a Feedforward Artificial Neural Network (FNN), specifically a Multilayer Perceptron (MLP) with fully connected. Instead of decision trees, ANN models utilized a structured network of neurons to learn patterns in the data.

For each selected feature subset, the dataset was split into training (70%) and testing (30%) sets. The ANN architecture consisted of an input layer matching the number of selected features, followed by two hidden layers with 64 and 32 neurons, respectively, using ReLU activation, and a single output neuron with linear activation for regression. The models were trained using the Adam optimizer with MSE loss, running for 100 epochs with a batch size of 32.

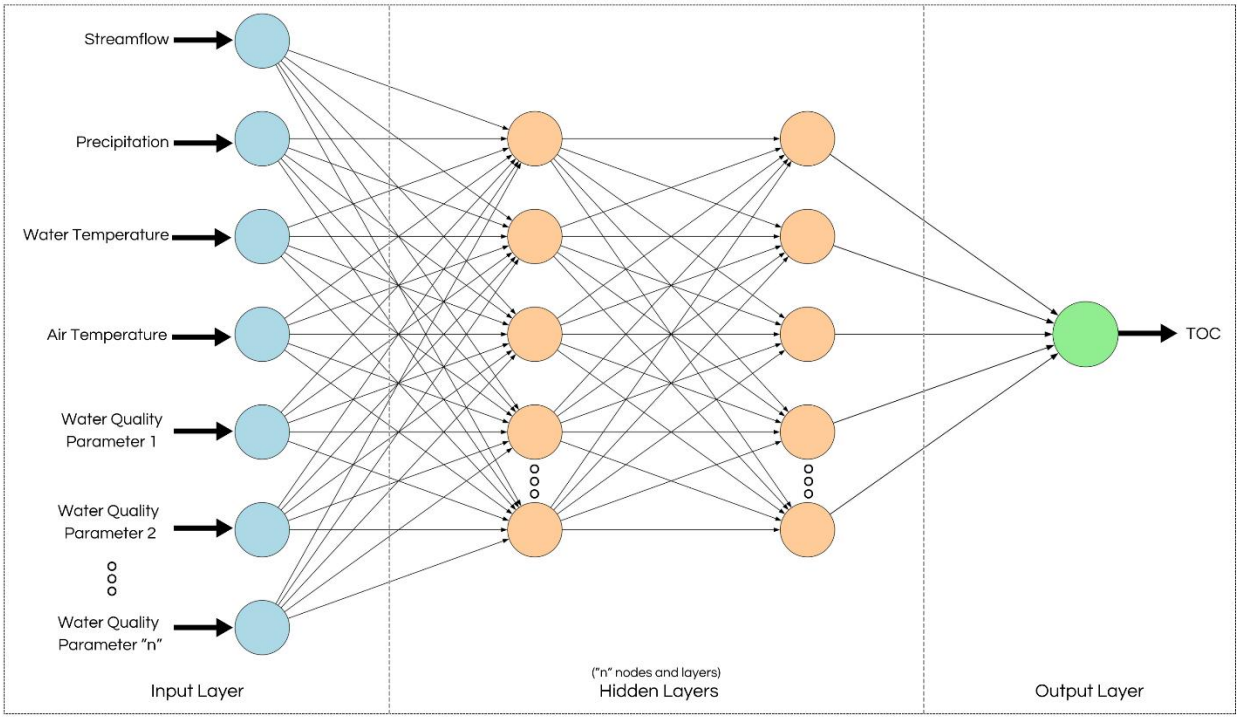

Figure 5. ANN Model Architecture

Performance evaluation followed the same approach as RF models, using MSE, MAE, and R^2 score, along with 5-fold cross-validation to assess generalization. The two best-performing feature subsets were selected for hyperparameter tuning, where Keras Tuner's Random Search was used to optimize the number of layers, neurons, activation functions, dropout rates, and learning rates. After tuning, the best models were retrained with optimized hyperparameters and evaluated using both train-test split and cross-validation metrics to confirm performance improvements.

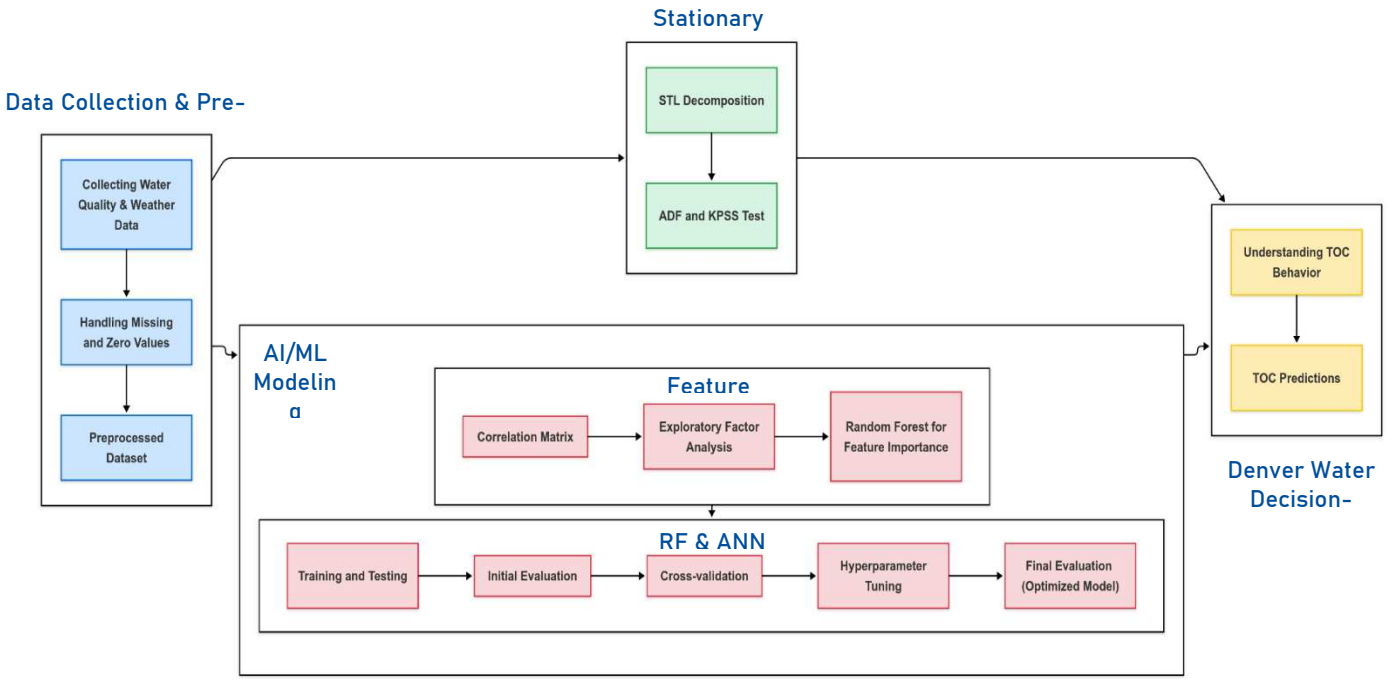

Figure 6. Research Methodology

CHAPTER 4 – RESULTS AND DISCUSSIONS

This chapter presents and discusses the findings from the stationarity analysis of TOC and evaluates the performance of the applied AI/ML models in predicting water quality.

4.1 Stationarity and non-stationarity of Total Organic Carbon

Figure 7 provides the results of the STL decomposition applied to the TOC time series across all sampling points, spanning roughly 15 years from 2010 to 2024. Results indicate that the historical TOC data or "Observed" component (first box, yellow line) fluctuate significantly between approximately 1 mg/L and 8 mg/L, with occasional sharp peaks. Notably, there are several higher-value peaks around 2013–2015 and again in 2023.

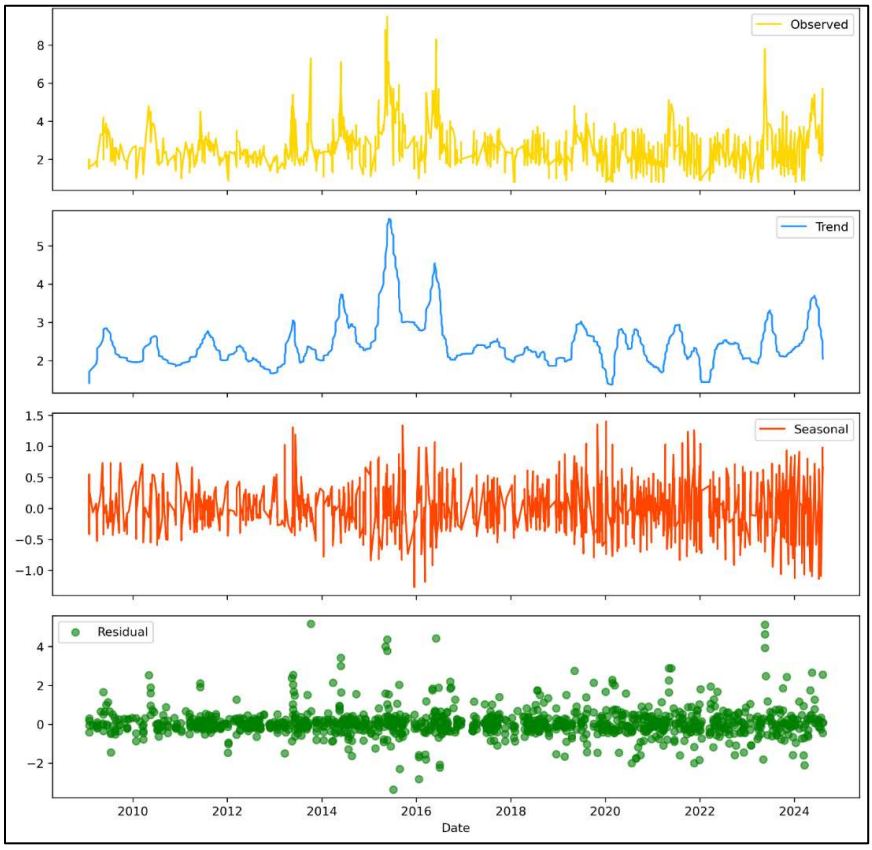

Figure 7. STL decomposition with LOESS for TOC

The elevated TOC levels from 2013–2015 may be linked to the persistent impacts of the 2012 “Lower North Fork” wildfire (Denver Water Geodatabase, 2025), as wildfire effects on water quality can persist for many years after the event. Meanwhile, the TOC peaks observed in 2020 appear to be associated with the “Cameron Peak” wildfire—Colorado’s largest wildfire of that year, burning 208,913 acres (National Interagency Coordination Center, 2020). Both wildfires had a notable impact on the overall water quality of the watershed, which has been corroborated by Denver Water experts (Personal, communication, 2025).

The second box (blue line) represents the “Trend” component. This smoothed line provides insight into the long-term behavior of the data by removing shorter-term fluctuations. The Trend generally increases from 2010 into the mid-2010s, reaching notable peaks around 2014–2015 and again around 2016. Both increasing trends of TOC could be attributed to the same wildfires mentioned in the “Observed” component.

The third panel (orange line) illustrates the “Seasonal” component, which reflects recurring cyclical patterns at set intervals, for example, differences in water quality between winter and summer. The amplitude of these oscillations ranges from around −1.5 to +1.5. A “negative” seasonal value means the TOC is below its normal level for that season. Conversely, a positive seasonal value indicates the TOC is above its typical seasonal level. Overall, the amplitude suggests a moderately sized seasonal effect that remains consistent throughout the time frame.

Table 4 provides the results of the ADF test. This test was designed to check for a unit root, with the null hypothesis being that the time series is non-stationary. In this case, the test statistic (approximately -5.85) is more negative than all the critical values, and the p-value (4.55×10^{-6}) is extremely small. Based on these results, we reject the null hypothesis and conclude that, according to the ADF test, the series is trend stationary. This means that once the trends mentioned above

are removed, the remaining time series should exhibit stationarity, that is, the fluctuations remain stable within a relatively consistent range over time, oscillating around a deterministic trend.

In contrast to ADF, the KPSS test shown in Table 4 uses a different null hypothesis where it states that the series is stationary or trend stationary. The test statistic (0.2710) is higher than the critical values at the 1% level (0.216), resulting in a p-value very small below 0.01. This leads us to reject null hypothesis in favor of non-stationarity by the KPSS test. There is not strong evidence that the series is non-stationary. It doesn't necessarily prove the series is stationary, but it indicates the test doesn't provide enough reason to doubt stationarity at typical confidence levels.

Table 4. Stationarity test for the TOC Time Series

Test	Statistic	p-value	1% Critical Value	5% Critical Value	10% Critical Value	Conclusion
ADF Test	-5.848199214	4.55E-06	-3.966905019	-3.414431338	-3.129370013	Reject H0: The series is trend-stationary.
KPSS Test	0.271090959	0.01	0.216	0.146	0.119	Reject H0: The series is NOT trend-stationary (non-stationary).

ADF and KPSS tests measure stationarity from opposite perspectives in their null hypothesis, the contradiction between the ADF and KPSS tests may stem from the fact that the ADF test primarily addresses stochastic trends but does not fully account for the seasonality identified in the data. Consequently, the ADF test could incorrectly suggest stationarity or trend-stationarity even though strong seasonal patterns persist that visible in the STL decomposition. Moreover, KPSS rejected trend stationarity because it sees seasonality as a trend that is not removed by simple detrending, KPSS account for seasonality that is usually considered a deterministic pattern because it repeats predictably over time.

Finally, the fourth panel (green points) depicts the "Residual" component, which represents the portion of the data left after removing both the Trend and Seasonal components, capturing the

unexplained variability in the time series, provides a detail of these residuals (Figure 8). These residuals are particularly important for this thesis because one of the research hypotheses is that they represent abnormal TOC patterns that can be used to test AI/ML models' ability to handle non-stationary data.

Most of the residual points cluster around zero, with a few sporadic outliers reaching as high as approximately 4 and as low as -2. The absence of clear patterns or systematic structure in the residuals suggests that the time series' behavior is primarily explained by the Trend and Seasonal components. As a result, the research hypothesis—suggesting that TOC exhibits non-stationary behavior due to an increase in extreme or abnormal weather events driven by climate change—is rejected. However, as noted, we are aware of two non-stationarity events that occurred during this period.

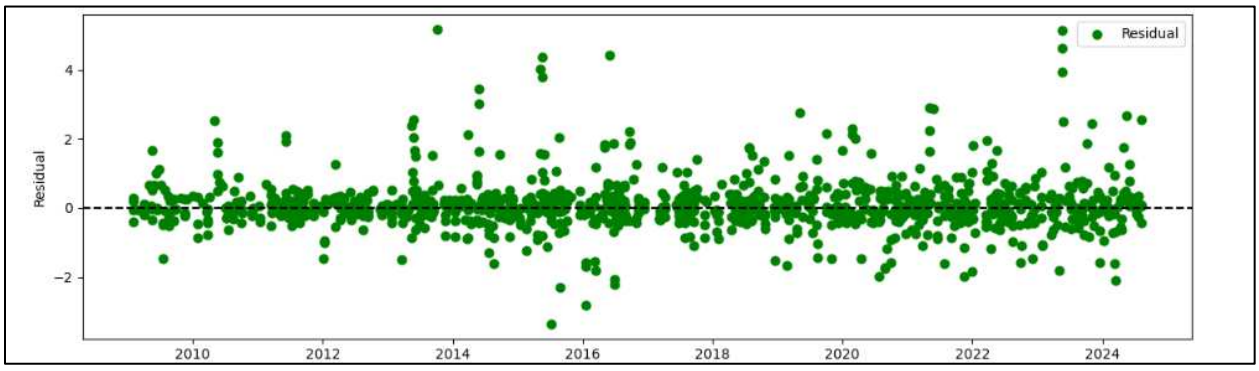

Figure 8. TOC Residuals

Table 5 lists the results of the ADF and KPSS tests on the residuals which provide additional support for stationary distribution. Specifically, the ADF test strongly indicates stationarity by rejecting its null hypothesis, suggesting that the statistical properties of the residuals remain stable over time.

Meanwhile, the KPSS test, with a p-value of 0.1, does not furnish enough evidence to reject the null hypothesis of stationarity. Therefore, the stationary analyses using STL, ADF, and KPSS collectively indicate that the TOC time series (and its residuals) is at least trend stationary.

Table 5. Stationarity test for the TOC residuals

Test	Statistic	p-value	1% Critical Value	5% Critical Value	10% Critical Value	Conclusion
ADF Test	-10.40390428	1.87839E-18	-3.436223123	-2.864133303	-2.56815075	Stationary (reject H0)
KPSS Test	0.024267876	0.1	0.739	0.463	0.347	Stationary (fail to reject H0)

Although the initial ADF and KPSS tests on the raw TOC time series yield conflicting conclusions—ADF suggesting trend-stationarity and KPSS indicating non-stationarity—STL decomposition shows that, once the main trend and seasonal components are removed, both tests agree the residuals are stationary. Consequently, the evidence supports treating the series as trend stationary. Although the series is trend-stationary, it is still non-stationary in its original form. As we do not remove the trend or seasonality in this study, the dataset used for modeling retains its non-stationary characteristics. The dataset has monthly frequency, so it has patterns of each seasonal phase and the observed trend progression to allow the models to capture these patterns. Additionally, since the residuals are found to be stationary and therefore do not indicate novel abnormal behaviors, this thesis did not analyze them separately; instead, they will be included alongside the complete time series without specific adjustments.

4.2 AI/ML modeling

This section discusses all results of the AI/ML modeling process, from feature selection to model training and testing.

4.2.1 Feature Selection

Different feature selection techniques were applied in this research to identify the most relevant input variables for the AI/ML models. This section provides a summary of results of these stages.

4.2.1.1 *Correlation Matrix*

The correlation matrix (Figure 9. **Correlation Matrix** provides an overview of the linear relationships among TOC and multiple water quality parameters, streamflow and weather data, offering insights into potential drivers of water chemistry variability. The strongest correlation observed was a negative association between TOC and sulfate (SO_4, r = -0.61), suggesting that organic matter degradation may drive sulfate reduction, particularly in anoxic or hypoxic conditions. This relationship is consistent with microbial sulfate reduction processes, where sulfate acts as an electron acceptor in the absence of oxygen, leading to its depletion in environments with high TOC.

TOC also shows significant positive correlations with flow-related parameters, such as the Chloride-to-Sulfate Mass Ratio (CSMR, r = 0.56) and instantaneous flow rate (Flow_inst, r = 0.51). This suggests that higher TOC concentrations may be associated with increased discharge or transport of organic material. The positive correlation with Total Suspended Solids (TSS, r = 0.46) further supports this trend, indicating that TOC is closely linked to particulate organic matter transport.

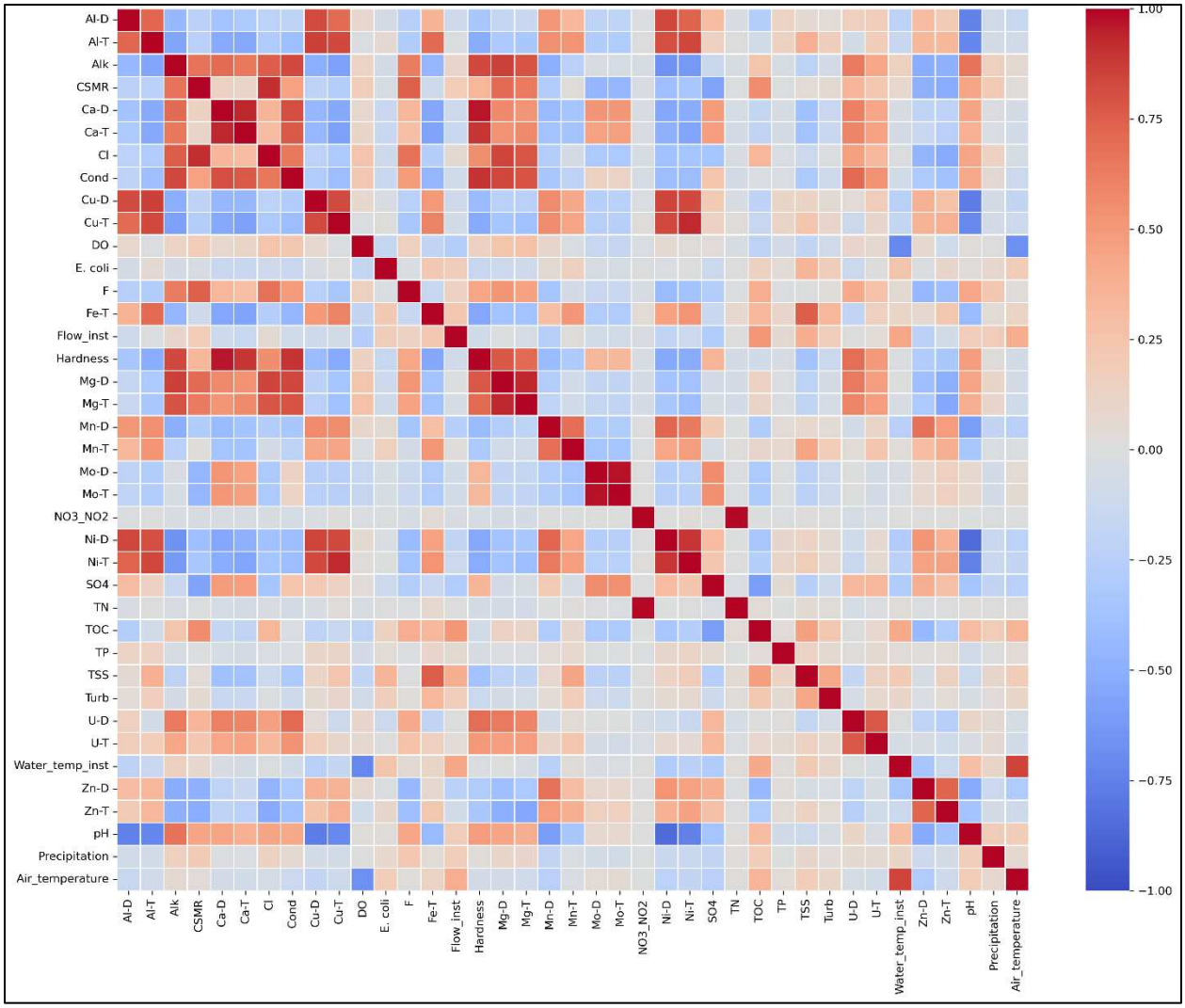

Figure 9. Correlation Matrix

Temperature-related variables were found to correlate with TOC variability, as seen in its positive correlations with water temperature ($r = 0.41$) and air temperature ($r = 0.35$). This suggests that TOC levels tend to be higher in warmer conditions, possibly due to enhanced microbial decomposition, algal growth, or increased leaching of organic material from soils and sediments. This temperature dependency highlights the potential influence of seasonal and climate-driven factors on organic carbon dynamics, with implications for long-term changes in water quality under global warming scenarios.

TOC exhibits moderate negative correlations with certain trace metals, including dissolved zinc (Zn-D, r = -0.44), nickel (Ni-D, r = -0.33), and molybdenum (Mo-D, r = -0.31). These negative relationships may indicate that TOC-rich waters facilitate the precipitation or adsorption of these metals.

4.2.1.2 *EFA*

Figure **10. Factor Loadings Heatmap** is the Factor Loadings Heatmap that provides information on how different water quality variables are associated with underlying latent factors extracted through EFA. Factor 1 appears to be dominated by water hardness and ionic strength, as evidenced by strong positive loadings on hardness (1.00), calcium (Ca-D: 0.97, Ca-T: 0.90), magnesium (Mg-D: 0.76, Mg-T: 0.69), conductivity (0.90), and chloride (0.53).

Factor 2 is characterized by sulfate-related geochemistry and potential industrial contamination, with high positive loadings on sulfate (SO_4: 0.59) and molybdenum (Mo-D: 0.72). The negative loading of the chloride-to-sulfate mass ratio (CSMR: -0.75) suggests that this factor represents waters dominated by sulfate rather than chloride.

Factor 3 is strongly associated with total and dissolved metal contamination, as well as suspended solids transport. High positive loadings are observed for total copper (Cu-T: 0.69), total nickel (Ni-T: 0.76), dissolved nickel (Ni-D: 0.81), total iron (Fe-T: 0.65), and total suspended solids (TSS: 0.66). This pattern suggests that Factor 3 represents metal pollution, likely influenced by sediment transport and erosion processes. These findings suggest that this factor could be linked to industrial or urban runoff, as well as natural weathering and erosion processes.

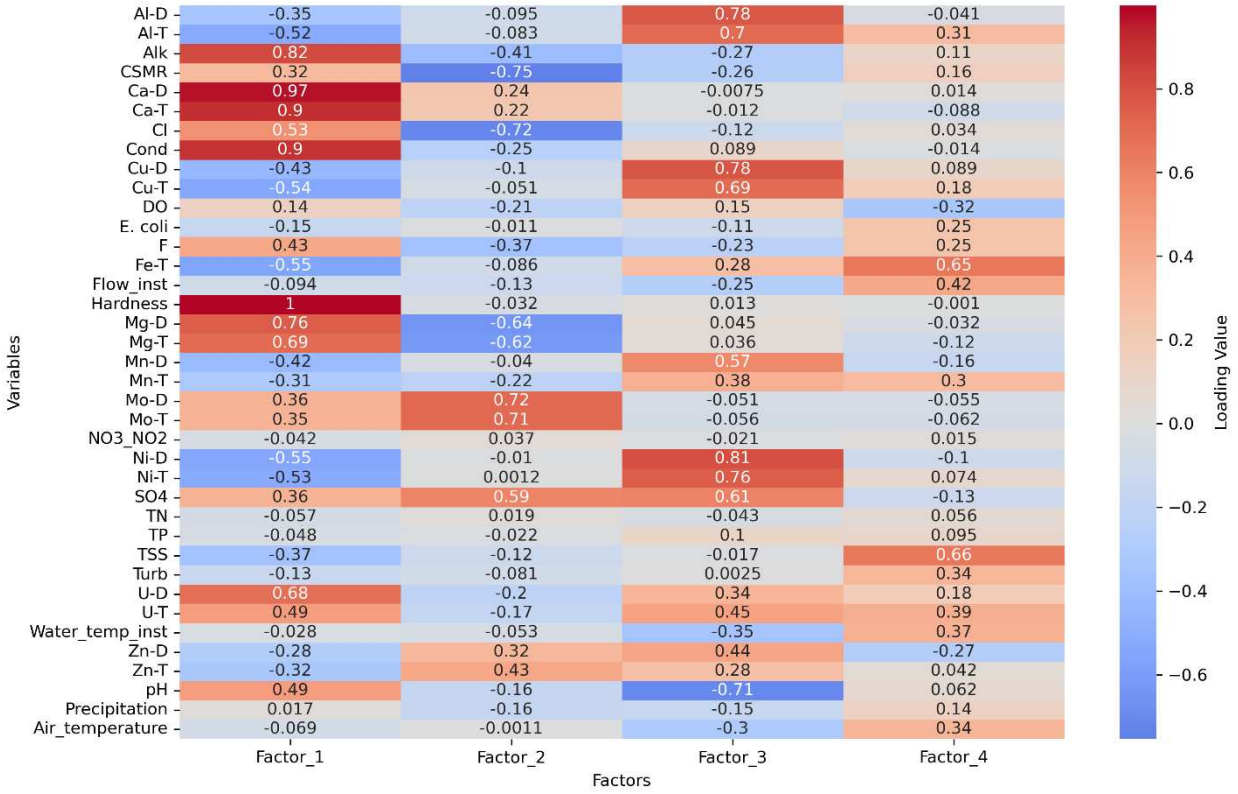

Figure 10. Factor Loadings Heatmap

Factor 4 appears related to hydrodynamic conditions, temperature, and sediment interactions. Positive loadings on flow rate (Flow_inst: 0.42), turbidity (0.39), and total iron (Fe-T: 0.65) suggest that this factor captures water movement and its role in sediment resuspension and metal transport. Interestingly, water temperature (Water_temp_inst: -0.35) loads negatively, suggesting that colder waters may be associated with higher concentrations of certain metals, possibly due to seasonal variations in solubility and precipitation. This factor highlights the importance of physical water movement in redistributing contaminants and influencing water quality.

4.2.1.3 *Feature Importance with Random Forest*

The feature importance analysis using Random Forest is summarized in **Figure 11**. This analysis provides valuable insights into the key variables influencing TOC in water systems. The

top 20 most important features have been ranked based on their contribution to TOC prediction, highlighting the role of geochemical, hydrological, and anthropogenic factors. The results indicate that TOC variability is driven by a combination of water chemistry, sediment transport, nutrient availability, and human influences.

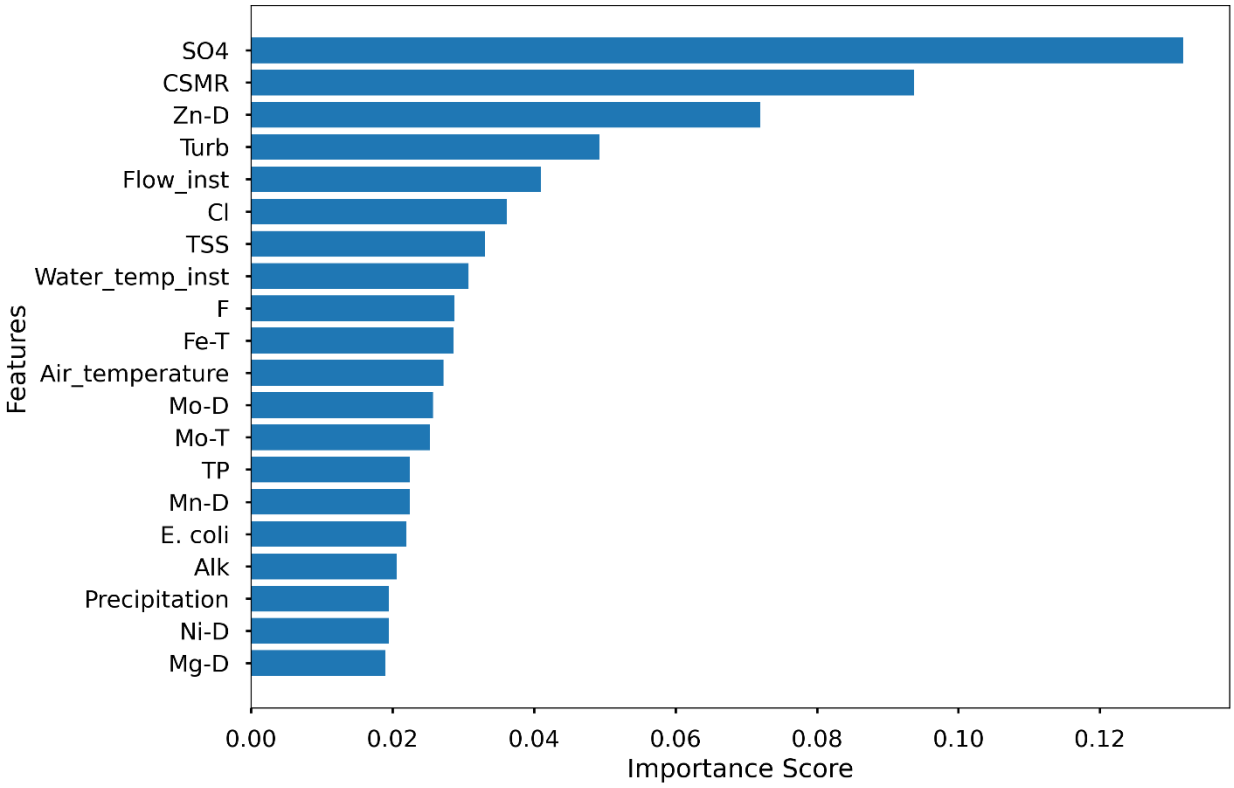

Figure 11. Top 20 Features by Importance (Random Forest Feature Selection)

Among all variables, sulfate (SO_4) emerges as the most important predictor of TOC, suggesting a strong connection between organic matter degradation and sulfate reduction processes. Another key factor is dissolved zinc (Zn-D). Turbidity (Turb) also shows a strong influence on TOC, reinforcing the idea that organic carbon is often associated with suspended particles in water.

Hydrological variables were also found to be a feature of TOC. Flow rate (Flow_inst) and total suspended solids (TSS) are among the most important features, suggesting that TOC is strongly influenced by water movement and sediment transport. Water temperature (Water_temp_inst) and air temperature (Air_temperature) further highlight the seasonal influence on TOC variability, with warmer conditions potentially enhancing organic matter decomposition and microbial activity.

A few additional variables, such as alkalinity (Alk), and precipitation, were also found to contribute to TOC variability but with relatively lower importance. Alkalinity (Alk) and precipitation indicate that carbonate buffering and rainfall-driven runoff may also play a role in TOC transport, albeit to a lesser extent.

4.2.1.4 Model selection

The optimal feature set for RF and ANN construction was selected based on the highest R^2_Mean_CV scores. These models were identified for both the complete dataset and the WS_LP_001 sampling point, ensuring that the best-performing configuration was chosen for further analysis. Once the best models were established, an additional assessment was conducted to evaluate the impact of including sampling point and sampling date as predictors on model performance. This evaluation aimed to determine how these categorical variables influenced the predictive capability of the model when included in different configurations.

The same subset of features was used across all predictor configurations, allowing for direct comparison of model performance under varying conditions. Specifically, three different configurations were tested: (1) using only the sampling point as a predictor, (2) using only the sampling date as a predictor, and (3) using both sampling point and sampling date as predictors. By systematically analyzing these variations, it was possible to assess the individual and combined

effects of these categorical variables on TOC prediction accuracy (Table 6 and Table 7). In all cases, TOC was excluded from the input variables.

Table 6. RF best models

Feature_Subset	R2_Mean_CV	Observations
Complete Dataset		
Top 1: Comprehensive Set (Top Features Consolidated)	0.7673	TOC, Sampling Point and Sampling Date were not included as predictors.
Top 1: Comprehensive Set (Top Features Consolidated)	0.7673	TOC and Sampling Date were not included as predictors.
Top 1: Baseline (All Features)	0.7665	TOC and Sampling Point were not included as predictors.
Sampling point WS_LP_001		
Top 1: Expanded Set 1 (Including Turbidity and Temperature) (with Flow_inst & Precipitation)	0.7423	TOC, Sampling Point and Sampling Date were not included as predictors.
Top 1: Expanded Set 1 (Including Turbidity and Temperature) (with Flow_inst & Precipitation)	0.7423	TOC and Sampling Date were not included as predictors.
Top 1: Expanded Set 1 (Including Turbidity and Temperature) (with Flow_inst & Precipitation)	0.7423	TOC and Sampling Point were not included as predictors.

Table 7. ANN best model

Feature_Subset	R2_Mean_CV	Observations
Complete Dataset		
Top1: Correlation_Features	0.7663	TOC, Sampling Point and Sampling Date were not included as predictors.
Top1: Correlation_Features	0.7783	TOC and Sampling Date were not included as predictors.
Top1: Correlation_Features	0.7546	TOC and Sampling Point were not included as predictors.
Sampling point WS_LP_001		
Top 1: Factor_Driven_Factor1	0.5037	TOC, Sampling Point and Sampling Date were not included as predictors.
Top 1: Factor_Driven_Factor2 (with Flow_inst & Precipitation)	0.5712	TOC and Sampling Date were not included as predictors.

Top 1: Factor_Driven_Factor2 (with Flow_inst & Precipitation)	0.6184	TOC and Sampling Point were not included as predictors.

4.2.2 Random Forest

The best-performing Random Forest models for the complete dataset utilized the "Comprehensive Set (Top Features Consolidated)" feature subset, achieving an R^2_Mean_CV of 0.7673 after hyperparameter tuning (Table 8). The "Top 10 Features (Random Forest Importance)" subset followed closely with R^2_Mean_CV = 0.7586.

During initial model evaluation, various feature selection techniques were tested to assess their impact on TOC prediction accuracy. The Baseline Model (All Features) performed well (R^2_Mean_CV = 0.7539), but smaller subsets, including the Top 20 Features (R^2_Mean_CV = 0.7531) and the Comprehensive Set (R^2_Mean_CV = 0.7588), achieved similar results. Additionally, the influence of precipitation and Flow_inst was examined. Incorporating these variables led to slight improvements, with the Comprehensive Set with Precipitation reaching R^2_Mean_CV = 0.7547 and the Top 10 Features with Precipitation achieving R^2_Mean_CV = 0.7493, reinforcing precipitation's role in TOC prediction.

Model stability, evaluated through cross-validation standard deviation (R^2_Std_CV), varied across feature sets. The Baseline Model (R^2_Std_CV = 0.0491) and Comprehensive Set with Precipitation (R^2_Std_CV = 0.0531) exhibited the lowest variability, indicating robust performance.

During final model evaluation (Table 9), slight performance improvements were observed from hyperparameter tuning, with the Comprehensive Set (Top Features Consolidated) achieving R^2_Mean_CV = 0.7673 (R^2_Std_CV = 0.0503) and the Top 10 Features (Random Forest Importance) yielding R^2_Mean_CV = 0.7586 (R^2_Std_CV = 0.0493).

Table 8. Initial RF Evaluation for the Complete Dataset

Feature Set	R2_Single_Split	MSE_Single_Split	MAE_Single_Split	R2_Mean_CV	R2_Std_CV	MSE_Mean_CV	MAE_Mean_CV
Baseline (All Features)	0.7885	0.2688	0.325	0.7539	0.0491	0.2941	0.3329
Correlation-Based Selection	0.7565	0.3095	0.342	0.7434	0.0517	0.3081	0.3425
Factor Analysis - Factor 1 (Hardness Components)	0.3695	0.8012	0.5531	0.4233	0.0528	0.707	0.5382
Factor Analysis - Factor 2 (Salinity and Metals)	0.5853	0.527	0.4565	0.5906	0.05	0.4999	0.4528
Factor Analysis - Factor 3 (Nickel and Copper)	0.631	0.4689	0.4393	0.6065	0.0548	0.4793	0.4492
Factor Analysis - Factor 4 (Suspended Solids)	0.2496	0.9536	0.7231	-0.0434	0.1071	1.2726	0.8272
Top 10 Features (Random Forest Importance)	0.7527	0.3142	0.3452	0.7569	0.048	0.2914	0.3412
Top 20 Features (Random Forest Importance)	0.7773	0.283	0.3293	0.7531	0.0487	0.296	0.3345
Overlapping Features (Across Methods)	0.7424	0.3274	0.3507	0.7027	0.0708	0.3531	0.3689
Expanded Set 1 (Including Turbidity and Temperature)	0.7307	0.3422	0.3547	0.7442	0.0498	0.3058	0.3502
Expanded Set 2 (Including Conductivity and Hardness)	0.7323	0.3401	0.3509	0.7514	0.048	0.2975	0.342
Comprehensive Set (Top Features Consolidated)	0.7561	0.31	0.3437	0.7588	0.0541	0.2873	0.335
Correlation-Based Selection (with Precipitation)	0.7582	0.3073	0.3421	0.7401	0.0509	0.3128	0.3466
Factor Analysis - Factor 1 (Hardness Components) (with Flow_inst & Precipitation)	0.5646	0.5532	0.4639	0.5362	0.0602	0.5626	0.4932
Factor Analysis - Factor 2 (Salinity and Metals) (with Flow_inst & Precipitation)	0.6508	0.4438	0.4174	0.6469	0.0566	0.4213	0.417
Factor Analysis - Factor 3 (Nickel and Copper) (with Flow_inst & Precipitation)	0.6694	0.4201	0.409	0.6409	0.0503	0.4329	0.4309
Factor Analysis - Factor 4 (Suspended Solids) (with Flow_inst & Precipitation)	0.4484	0.7009	0.5949	0.3907	0.1463	0.7296	0.6131
Top 10 Features (Random Forest Importance) (with Precipitation)	0.7562	0.3099	0.3446	0.7493	0.0509	0.2991	0.3451
Overlapping Features (Across Methods) (with Flow_inst & Precipitation)	0.7573	0.3084	0.3389	0.7087	0.0729	0.348	0.3613
Expanded Set 1 (Including Turbidity and Temperature) (with Flow_inst & Precipitation)	0.7548	0.3115	0.3407	0.7458	0.0602	0.3023	0.3455
Expanded Set 2 (Including Conductivity and Hardness) (with Flow_inst & Precipitation)	0.7529	0.3141	0.3447	0.753	0.0542	0.2945	0.3413
Comprehensive Set (Top Features Consolidated) (with Precipitation)	0.7638	0.3002	0.3397	0.7547	0.0531	0.2929	0.3393

Table 9. Final RF Evaluation for the Complete Dataset After Tuning

Feature Set	Best Parameters	R^2 (Single Split)	R^2 (Mean CV)	R^2 (Std CV)
Comprehensive Set (Top Features Consolidated)	{'bootstrap': True, 'max_depth': 20, 'max_features': 'sqrt', 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 200}	0.7671	0.7673	0.0503
Top 10 Features (Random Forest Importance)	{'bootstrap': True, 'max_depth': 30, 'max_features': None, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 500}	0.7532	0.7586	0.0493

For the WS_LP_001 sampling point, the best-performing Random Forest (RF) models during both the initial and final evaluations utilized the "Expanded Set 1 (Including Turbidity and Temperature) with Flow_inst & Precipitation" feature subset, achieving an R^2_Mean_CV = 0.7423 with R^2_Std_CV = 0.0792. Similarly, the "Top 10 Features (Random Forest Importance) with Precipitation" subset performed comparably, yielding an R^2_Mean_CV = 0.7432 with a lower R^2_Std_CV of 0.0496 (Table 10. **Final RF Evaluation for sampling point WS-LP-001** and Table 11).

It was noted that the R^2 values from a single train-test split (R^2_Single_Split) for these models were lower than their cross-validation (CV) scores. The "Expanded Set 1 (Including Turbidity and Temperature) with Flow_inst & Precipitation" subset achieved an R^2_Single_Split = 0.5822, while the "Top 10 Features (Random Forest Importance) with Precipitation" subset yielded an R^2_Single_Split = 0.5928. The discrepancy between the R^2_Mean_CV and R^2_Single_Split is likely due to the small dataset size at WS_LP_001 compared to the complete dataset. With the limited dataset this thesis was not fully able to determine whether a single train-test split may provide a representation of all conditions.

Table 10. Final RF Evaluation for sampling point WS-LP-001

Feature Set	Best Parameters	R^2 (Single Split)	R^2 (mean CV)	R^2 (Std CV)
Expanded Set 1 (Including Turbidity and Temperature) (with Flow_inst & Precipitation)	{'bootstrap': True, 'max_depth': 10, 'max_features': None, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 200}	0.5822	0.7423	0.0792
Top 10 Features (Random Forest Importance) (with Precipitation)	{'bootstrap': True, 'max_depth': 10, 'max_features': 'sqrt', 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 100}	0.5928	0.7432	0.0496

Table 11. Initial RF Evaluation for sampling point WS-LP-001

Feature Set	R2_Single_Split	MSE_Single_Split	MAE_Single_Split	R2_Mean_CV	R2_Std_CV	MSE_Mean_CV	MAE_Mean_CV
Baseline (All Features)	0.6092	0.2632	0.3308	0.6949	0.0823	0.1886	0.3167
Correlation-Based Selection	0.5676	0.2912	0.352	0.7089	0.0904	0.1823	0.3157
Factor Analysis - Factor 1 (Hardness Components)	0.4958	0.3395	0.379	0.4971	0.0957	0.3219	0.4106
Factor Analysis - Factor 2 (Salinity and Metals)	0.4516	0.3693	0.4177	0.4275	0.1877	0.3527	0.4248
Factor Analysis - Factor 3 (Nickel and Copper)	0.533	0.3145	0.4124	0.5917	0.1047	0.2558	0.3635
Factor Analysis - Factor 4 (Suspended Solids)	0.0693	0.6268	0.5847	0.072	0.3698	0.6054	0.5808
Top 10 Features (Random Forest Importance)	0.5931	0.274	0.3543	0.732	0.0667	0.1657	0.3014
Top 20 Features (Random Forest Importance)	0.5972	0.2713	0.3463	0.726	0.085	0.1652	0.2968
Overlapping Features (Across Methods)	0.5511	0.3023	0.3729	0.5128	0.1335	0.3103	0.3746
Expanded Set 1 (Including Turbidity and Temperature)	0.5537	0.3006	0.3784	0.6573	0.1145	0.2046	0.3246
Expanded Set 2 (Including Conductivity and Hardness)	0.5948	0.2729	0.3516	0.6518	0.1234	0.2065	0.3278
Comprehensive Set (Top Features Consolidated)	0.5814	0.2819	0.3559	0.7273	0.0758	0.1663	0.308
Correlation-Based Selection (with Precipitation)	0.5616	0.2952	0.3532	0.7042	0.1024	0.1835	0.3134
Factor Analysis - Factor 1 (Hardness Components) (with Flow_inst & Precipitation)	0.4602	0.3635	0.3973	0.6284	0.1257	0.2358	0.3463
Factor Analysis - Factor 2 (Salinity and Metals) (with Flow_inst & Precipitation)	0.5154	0.3264	0.3912	0.6913	0.1094	0.1873	0.3271
Factor Analysis - Factor 3 (Nickel and Copper) (with Flow_inst & Precipitation)	0.5401	0.3097	0.3854	0.6921	0.1034	0.187	0.3201
Factor Analysis - Factor 4 (Suspended Solids) (with Flow_inst & Precipitation)	0.2779	0.4863	0.4978	0.3933	0.1984	0.381	0.4635
Top 10 Features (Random Forest Importance) (with Precipitation)	0.5998	0.2695	0.3512	0.7334	0.0711	0.1644	0.2997
Overlapping Features (Across Methods) (with Flow_inst & Precipitation)	0.5554	0.2995	0.3561	0.7	0.1172	0.1856	0.3119
Expanded Set 1 (Including Turbidity and Temperature) (with Flow_inst & Precipitation)	0.5908	0.2756	0.359	0.7373	0.074	0.1619	0.3024
Expanded Set 2 (Including Conductivity and Hardness) (with Flow_inst & Precipitation)	0.5869	0.2782	0.3419	0.7162	0.0797	0.1741	0.3103
Comprehensive Set (Top Features Consolidated) (with Precipitation)	0.5832	0.2807	0.3545	0.7274	0.0717	0.1674	0.3062

4.2.3 ANN

The ANN evaluation for the complete dataset is given in Table 12. The best-performing ANN model, based on R^2_Mean_CV, utilized the "Comprehensive Set (Top Features Consolidated) (with Precipitation)" feature subset, achieving an R^2_Mean_CV = 0.7611 after hyperparameter tuning. This was followed by the "Comprehensive Set (Top Features Consolidated)" subset, which achieved an R^2_Mean_CV = 0.7112 (Table 12 and Table 13).

These results demonstrate the impact of feature selection techniques in improving model performance. The inclusion of precipitation in the feature set contributed to better generalization, as indicated by the higher R^2_Mean_CV and improved stability (lower R^2_Std_CV = 0.0315 compared to 0.0618 for the model without precipitation).

Table 12. Initial ANN Evaluation for the Complete Dataset

Feature_Set	R2_Single_Split	MSE_Single_Split	MAE_Single_Split	R2_Mean_CV	R2_Std_CV
Comprehensive Set (Top Features Consolidated) (with Precipitation)	0.6892	0.395	0.4048	0.7524	0.0256
Comprehensive Set (Top Features Consolidated)	0.6824	0.4035	0.407	0.7498	0.0272
Correlation-Based Selection	0.7265	0.3476	0.3755	0.7493	0.039
Overlapping Features (Across Methods) (with Flow_inst & Precipitation)	0.6948	0.3879	0.3993	0.7483	0.0278
Correlation-Based Selection (with Precipitation)	0.6952	0.3874	0.3933	0.7482	0.0168
Overlapping Features (Across Methods)	0.6585	0.434	0.4116	0.7436	0.0434
Top 20 Features (Random Forest Importance)	0.6247	0.4769	0.3989	0.7351	0.0267
Expanded Set 1 (Including Turbidity and Temperature)	0.6866	0.3982	0.3907	0.7349	0.0476
Expanded Set 1 (Including Turbidity and Temperature) (with Flow_inst & Precipitation)	0.6046	0.5025	0.4223	0.7308	0.0464
Expanded Set 2 (Including Conductivity and Hardness)	0.7558	0.3103	0.3802	0.729	0.0442
Expanded Set 2 (Including Conductivity and Hardness) (with Flow_inst & Precipitation)	0.7006	0.3804	0.3962	0.7279	0.0517
Top 10 Features (Random Forest Importance) (with Precipitation)	0.6549	0.4385	0.3991	0.7199	0.0397
Top 10 Features (Random Forest Importance)	0.6485	0.4467	0.4028	0.7156	0.0699
Factor Analysis - Factor 2 (Salinity and Metals) (with Flow_inst & Precipitation)	0.6632	0.428	0.439	0.694	0.0333
Baseline (All Features)	0.439	0.713	0.4545	0.6811	0.0526
Factor Analysis - Factor 3 (Nickel and Copper) (with Flow_inst & Precipitation)	0.602	0.5058	0.4742	0.647	0.072
Factor Analysis - Factor 3 (Nickel and Copper)	0.5407	0.5837	0.5041	0.6144	0.0713
Factor Analysis - Factor 2 (Salinity and Metals)	0.6136	0.4911	0.4708	0.6137	0.055
Factor Analysis - Factor 1 (Hardness Components) (with Flow_inst & Precipitation)	0.4934	0.6437	0.5223	0.5224	0.0629
Factor Analysis - Factor 1 (Hardness Components)	0.4074	0.753	0.5522	0.4355	0.0672
Factor Analysis - Factor 4 (Suspended Solids) (with Flow_inst & Precipitation)	0.4351	0.7178	0.6121	0.3713	0.0383
Factor Analysis - Factor 4 (Suspended Solids)	0.3588	0.8148	0.6765	0.2619	0.1004

Table 13. Final ANN Evaluation for the Complete Dataset

Feature Set	Best Parameters	R^2 (Single Split)	R^2 (mean CV)	R^2 (Std CV)
Comprehensive Set (Top Features Consolidated) (with Precipitation)	{'num_layers': 4, 'units': [256, 32, 64, 64], 'activation': 'tanh', 'dropout_rate': 0.1, 'learning_rate': 0.0010028966014693037}	0.7429	0.7611	0.0315
Comprehensive Set (Top Features Consolidated)	{'num_layers': 3, 'units': [96, 64, 96], 'activation': 'tanh', 'dropout_rate': 0.0, 'learning_rate': 0.0007358058711559769}	0.7403	0.7112	0.0618

For the WS_LP_001 sampling point, the best-performing ANN models during both the initial and final evaluations utilized the "Expanded Set 1 (Including Turbidity and Temperature) with Flow_inst & Precipitation" feature subset, achieving an R^2_Mean_CV = 0.7423 with R^2_Std_CV = 0.0792. Similarly, the "Top 10 Features (Random Forest Importance) with Precipitation" subset performed comparably, yielding an R^2_Mean_CV = 0.7432 with a lower R^2_Std_CV of 0.0496 (Table 144. **Initial ANN Evaluation for sampling point WS-LP-001** and Table 155. **Final ANN Evaluation for sampling point WS-LP-001**). However, as the R^2 values from a single train-test split (R^2_Single_Split) for these models were significantly lower than their cross-validation (CV) scores. The "Expanded Set 1 (Including Turbidity and Temperature) with Flow_inst & Precipitation" subset achieved an R^2_Single_Split = 0.5822, while the "Top 10 Features (Random Forest Importance) with Precipitation" subset yielded an R^2_Single_Split = 0.5928. As with RF models, the discrepancy between the R^2_Mean_CV and R^2_Single_Split is likely due to the small dataset size at WS_LP_001 compared to the complete dataset.

With a limited dataset, a single train-test split may not provide the representation on non-stationarity this thesis was investigating. For example, TOC may be unevenly distributed in a single test set. This imbalance may have led to unrepresentative test conditions.

Table 144. Initial ANN Evaluation for sampling point WS-LP-001

Feature_Set	R2_Single_Split	MSE_Single_Split	MAE_Single_Split	R2_Mean_CV	R2_Std_CV
Factor Analysis - Factor 2 (Salinity and Metals) (with Flow_inst & Precipitation)	0.2874	0.4799	0.4546	0.5114	0.1692
Factor Analysis - Factor 1 (Hardness Components)	0.4639	0.3611	0.4177	0.4989	0.0819
Factor Analysis - Factor 2 (Salinity and Metals)	0.4761	0.3528	0.4218	0.4584	0.2131
Factor Analysis - Factor 1 (Hardness Components) (with Flow_inst & Precipitation)	0.3897	0.411	0.4612	0.4551	0.1324
Top 20 Features (Random Forest Importance)	0.2213	0.5245	0.5164	0.4253	0.2216
Top 10 Features (Random Forest Importance) (with Precipitation)	-0.0212	0.6877	0.6105	0.4013	0.1632
Factor Analysis - Factor 4 (Suspended Solids) (with Flow_inst & Precipitation)	-0.1783	0.7935	0.615	0.3688	0.237
Top 10 Features (Random Forest Importance)	0.3603	0.4308	0.507	0.3479	0.2438
Expanded Set 1 (Including Turbidity and Temperature) (with Flow_inst & Precipitation)	-0.4745	0.993	0.6312	0.1381	0.674
Baseline (All Features)	0.0124	0.6651	0.5894	0.0885	0.4295
Correlation-Based Selection (with Precipitation)	-0.3922	0.9376	0.5847	0.0723	0.9133
Factor Analysis - Factor 4 (Suspended Solids)	-1.3575	1.5877	0.7334	0.0138	0.3796
Comprehensive Set (Top Features Consolidated) (with Precipitation)	-0.433	0.965	0.5866	0.0016	0.9162
Correlation-Based Selection	-0.0937	0.7365	0.5389	-0.2311	1.4062
Comprehensive Set (Top Features Consolidated)	-1.0404	1.3741	0.6554	-0.2733	1.6331
Expanded Set 2 (Including Conductivity and Hardness)	-1.4353	1.6401	0.6131	-0.3489	1.6707
Expanded Set 2 (Including Conductivity and Hardness) (with Flow_inst & Precipitation)	-0.5987	1.0767	0.6057	-0.3627	1.7901
Overlapping Features (Across Methods) (with Flow_inst & Precipitation)	-0.7815	1.1998	0.5819	-0.3797	1.7858
Factor Analysis - Factor 3 (Nickel and Copper) (with Flow_inst & Precipitation)	-1.7566	1.8565	0.6939	-0.3869	0.4327
Overlapping Features (Across Methods)	-2.1518	2.1226	0.6564	-0.415	1.7797
Factor Analysis - Factor 3 (Nickel and Copper)	-1.9318	1.9745	0.6761	-0.5441	0.769
Expanded Set 1 (Including Turbidity and Temperature)	-2.1231	2.1033	0.6982	-0.6326	2.096

Table 155. Final ANN Evaluation for sampling point WS-LP-001

Feature Set	Best Parameters	R^2 (Single Split)	R^2 (mean CV)	R^2 (Std CV)
Factor Analysis - Factor 2 (Salinity and Metals) (with Flow_inst & Precipitation)	{'num_layers': 3, 'units': [256, 96, 192], 'activation': 'relu', 'dropout_rate': 0.0, 'learning_rate': 0.0023479211967056767}	0.4097	0.4937	0.168
Factor Analysis - Factor 1 (Hardness Components)	{'num_layers': 4, 'units': [256, 128, 128, 64], 'activation': 'tanh', 'dropout_rate': 0.4, 'learning_rate': 0.006831449750329013}	0.4272	0.4919	0.1628

For the WS_LP_001 sampling point during the initial evaluation, the best-performing ANN model utilized "Factor Analysis - Factor 2 (Salinity and Metals) (with Flow_inst & Precipitation)," achieving an R^2_Mean_CV = 0.5114. This was closely followed by "Factor Analysis - Factor 1 (Hardness Components)" with an R^2_Mean_CV = 0.4989, suggesting that salinity, metals, and hardness-related features played a crucial role in predictive performance for this specific sampling point.

After hyperparameter tuning, the best-performing model remained the same, with "Factor Analysis - Factor 2 (Salinity and Metals) (with Flow_inst & Precipitation)" achieving an R^2_Mean_CV = 0.4937. Similarly, "Factor Analysis - Factor 1 (Hardness Components)" maintained its strong performance with an R^2_Mean_CV = 0.4919, reinforcing the importance of these feature subsets. While the R^2 scores slightly decreased after tuning (from 0.5114 to 0.4937 for Factor Analysis - Factor 2 and from 0.4989 to 0.4919 for Factor Analysis - Factor 1), the hyperparameter optimization process significantly improved model stability by reducing performance variance across different training splits. This suggests that although the models exhibit a minor reduction in predictive power, they are now less prone to overfitting and more generalizable, potentially making these models better suited for our area and other applications where they will encounter unseen data.

4.2.4 Model Comparison

The performance comparison between RF and ANN (Table 16) demonstrates that both models achieved high R^2 scores and low standard deviations, benefiting from the larger dataset size in the complete dataset. This indicates that both RF and ANN effectively capture the underlying patterns when sufficient data is available.

The best-performing RF model (Comprehensive Set - Top Features Consolidated) achieved an R^2_Mean_CV of 0.7673, which is closely comparable to the best ANN model (Comprehensive Set with Precipitation), which achieved an R^2_Mean_CV of 0.7611. The small difference between these R^2 values suggests that both models are equally effective in learning the relationships within the dataset.

Additionally, the Comprehensive Set consistently outperformed the Top 10 Features in both RF and ANN, highlighting the importance of using a well-balanced and broader feature selection approach. However, the Top 10 Features (RF-based selection) still achieved a strong R^2_Mean_CV of 0.7586, demonstrating that a more compact feature set can maintain high predictive performance.

Table 16. Performance comparison between RF and ANN for the complete dataset

Feature Set	Model	R^2 (Single Split)	R^2 (Mean CV)	R^2 (Std CV)
Comprehensive Set (Top Features Consolidated)	RF	0.7671	0.7673	0.0503
Top 10 Features (Random Forest Importance)	RF	0.7532	0.7586	0.0493
Comprehensive Set (Top Features Consolidated) (with Precipitation)	ANN	0.7429	0.7611	0.0315
Comprehensive Set (Top Features Consolidated)	ANN	0.7403	0.7112	0.0618

For sampling point WS-LP-001 (Table 17), a significant drop in R^2 scores from CV to Single Split was observed, indicating potential generalization issues due to the limited dataset size. RF models exhibited substantial drops in performance, suggesting overfitting. For Expanded Set

1 (RF), R^2_Mean_CV dropped from 0.7423 to 0.5822 (Δ = -0.1601), while for the Top 10 Features (RF), it decreased from 0.7432 to 0.5928 (Δ = -0.1504). ANN models, although performing worse overall, experienced a smaller decline. For Factor Analysis - Factor 2 (Salinity and Metals) (ANN), R^2_Mean_CV dropped from 0.4937 to 0.4097 (Δ = -0.0840), and for Factor Analysis - Factor 1 (Hardness Components) (ANN), it decreased from 0.4919 to 0.4272 (Δ = -0.0647).

The large drop in R^2 for RF models suggests that they performed well in cross-validation but struggled to maintain their performance when applied to a new test set, indicating overfitting due to the limited dataset size. This suggests that while the models captured patterns well within the training folds, they failed to generalize effectively to unseen data. In contrast, ANN models exhibited a smaller decline in performance from CV to Single Split, indicating that while they struggled with overall predictive power, they were less prone to overfitting than RF models.

ANN models are generally more sensitive to data size, requiring larger datasets to fully capture meaningful patterns. The limited sample size at WS-LP-001 negatively impacted their performance, as evidenced by their lower R^2 scores and higher standard deviation (R^2_Std_CV). RF models, on the other hand, are typically more robust in low-sample, high-dimensional problems, making them a more suitable choice for smaller datasets. However, given the significant drop in R^2 and generalization issues, neither RF nor ANN provides sufficient predictive accuracy for real-world TOC evaluation at this location.

Table 17. Performance comparison between RF and ANN for sampling point WS-LP-001

Feature Set	Model	R^2 (Single Split)	R^2 (Mean CV)	R^2 (Std CV)
Expanded Set 1 (Including Turbidity and Temperature) (with Flow_inst & Precipitation)	RF	0.5822	0.7423	0.0792
Top 10 Features (Random Forest Importance) (with Precipitation)	RF	0.5928	0.7432	0.0496
Factor Analysis - Factor 2 (Salinity and Metals) (with Flow_inst & Precipitation)	ANN	0.4097	0.4937	0.168
Factor Analysis - Factor 1 (Hardness Components)	ANN	0.4272	0.4919	0.1628

CHAPTER 5 – CONCLUSIONS

This thesis explored non-stationarity of water quality using TOC as a proxy. Water quality data from the Denver metro area included water quality peaks between 2013–2015 and 2023. These were known abnormalities. Stationarity test produced conflicting results and residual analysis suggested TOC time series is primarily trend stationary. A secondary analysis of AI RF and ANN models to predict TOC resulted in positive outcomes. While RF and ANN models performed (lower errors) well on the complete dataset, their performance declined when predicting on a single sampling point. In predicting water quality and informing more non-stationarity knowledge, this thesis found that more data is essential. Additional frequency of data collection and integrating weather stations would improve model accuracy and possible provide reliable early warning system for Denver Water Treatment Operations. Additionally, further stationarity analysis and testing STL residuals in AI models could provide insights of the importance of the residuals and how impact AI model performances. Additional research observations include:

- The results of the stationarity tests ADF and KPSS tests contradict each other. The ADF test indicates that the TOC time series exhibits trend-stationary behavior, while the KPSS test suggests non-stationarity. The ADF test confirms stationarity in the residuals, while the KPSS test does not provide strong evidence to reject stationarity at typical confidence levels, therefore the TOC time series is trend stationary. These differences could be due to the ADF test is more sensitive to stochastic trends, while KPSS is stricter and treats seasonality as a trend.
- According to the RF feature importance ranking, SO_4 emerges as the most significant predictor of TOC. River flow and TSS are also among the most important features,

confirming research that TOC is heavily influenced by water movement and sediment transport.

- The performance comparison between RF and ANN shows that both models achieved high R^2 scores and low standard deviations. Suggesting that both RF and ANN are effective predictive tools across all sampling points. However as both models saw performance decline in sampling point WS-LP-001 estimations, more research is needed to determine which model would provide sufficient decision support for water operations.
- Low correlation with weather suggests misalignment (temporally and spatially) with TOC data. Given the sparsity of water quality data, it is likely that more frequent water quality samples would see more correlation. However, more weather stations at key water quality stations would also provide a more comprehensive and useful modeling dataset. This would enable water quality data at each sampling point to be directly associated with specific local weather data.
- A more extensive analysis of stationary water quality behavior should be conducted of residual data. Combined residual analysis and AI model predictions could improve the understanding of non-stationarity events as new abnormal patterns develop.

5.1 Implications of Results

This thesis investigated how AI/ML can assist water utilities in managing non-stationary water quality states. If found RF and ANN models had a high predictive capacity. It also found that residuals, data left after removing both the trend and seasonal patterns, could provide a means for detecting non-stationarity. If the residuals were non-stationary, they would highlight abnormal TOC patterns that AI/ML models could learn. However, conflicting results from ADF and KPSS

tests indicate that the test dataset was likely insufficient to verify this premise. The implications and contributions of this research are discussed in the following sections.

5.2 Contributions and implications of the research

This research has been contributed to body of knowledge in the following aspects:

- The developed System Thinking Map illustrates the interconnected effects of climate change, extreme weather events, and water quality dynamics on drinking water treatment. It also highlights how the integration of an effective AI system can reduce operational uncertainties through anomaly detection and predictive capabilities, ultimately enhancing decision-making and system resilience.
- Investigating how ADF and KPSS tests provided statistical evidence to determine whether a time series is stationary or not.
- Exploring how AI/ML techniques can enhance Denver Water operations by developing a TOC predictive tool and assessing their capability to handle non-stationary patterns through the integration of TOC residual analysis.

5.3 Limitations and Challenges

An optimized AI model that accurately predicts TOC levels at individual sampling points, such as those that serve as an influent proxy for water treatment plants, would provide decision support for plant operations. However, the main limitation of this thesis was the lack of sufficient water quality data, as the available data is collected only at a monthly frequency. This limited data constrains the ability of AI/ML models to predict short term fluctuations and dynamic patterns. Additionally, the lack of finding non-stationarity in the residual data (as identified through STL decomposition) prevented testing whether RF and ANN models could effectively handle non-stationary patterns.

5.4 Future research directions

Further research should be undertaken to explore the use of AI/ML models in addressing non-stationarity water quality changes. First, additional non-stationary detection techniques should be applied beyond ADF and KPSS. Second, the models developed in this research should be applied to new, unseen data from Denver Water to reassess their performance and validate their reliability. Third, further residuals analysis should be undertaken as input to AI models to explore residuals exhibit non-stationary behavior. Fourth, other AI models should be explored. For instance, how might reinforcement learning models support adaptive models that dynamically respond to water quality changes. Fifth, shifting the focus from watershed conditions to the water quality of the influent entering the Foothills Treatment Plant might result in different outcomes.

REFERENCES

1. Ahmad, M., Ahmad, A., Omar, T. F. T., & Mohammad, R. (n.d.). Current Trends of Analytical Techniques for Total Alkalinity Measurement in Water Samples: A Review. Critical Reviews in Analytical Chemistry, 0(0), 1–11. https://doi.org/10.1080/10408347.2023.2199432
2. Albanakis, C., Tsanana, E., & Fragkaki, A. G. (2021). Modeling and prediction of trihalomethanes in the drinking water treatment plant of Thessaloniki, Greece. Journal of Water Process Engineering, 43, 102252. https://doi.org/10.1016/j.jwpe.2021.102252
3. Alnahit, A. O., Mishra, A. K., & Khan, A. A. (2022). Stream water quality prediction using boosted regression tree and random forest models. Stochastic Environmental Research and Risk Assessment, 36(9), 2661–2680. https://doi.org/10.1007/s00477-021-02152-4
4. Al-Obaidi, B. H. K., Ali, S. K., & Jassim, D. T. (2020). Influence of a river water quality on the efficiency of water treatment using artificial neural network. Journal of Engineering Science and Technology, 15(4), 2610-2623.
5. Alomani, S. M., Alhawiti, N. I., & Alhakamy, A. (2022). Prediction of Quality of Water According to a Random Forest Classifier. International Journal of Advanced Computer Science and Applications, 13(6). https://doi.org/10.14569/IJACSA.2022.01306105
6. Assefa, E., Jabasingh, A., Tadesse, A. M., Dessalegne, M., Mulugeta, E., & Teju, E. (2024). Seasonal natural removal and implications for disinfection byproduct formation at the Koka water treatment plant: Upper Awash, Ethiopia. Journal of Water and Health, jwh2024230. https://doi.org/10.2166/wh.2024.230
7. Behmel, S., Damour, M., Ludwig, R., & Rodriguez, M. J. (2016). Water quality monitoring strategies—A review and future perspectives. Science of The Total Environment, 571, 1312–1329. https://doi.org/10.1016/j.scitotenv.2016.06.235
8. Bertone, E., Sahin, O., Richards, R., & Roiko, A. (2016). Extreme events, water quality and health: A participatory Bayesian risk assessment tool for managers of reservoirs. Journal of Cleaner Production, 135, 657–667. https://doi.org/10.1016/j.jclepro.2016.06.158
9. Blackburn, E. A. J., Dickson-Anderson, S. E., Anderson, W. B., & Emelko, M. B. (2023). Biological Filtration is Resilient to Wildfire Ash-Associated Organic Carbon Threats to Drinking Water Treatment. ACS ES&T Water, 3(3), 639–649. https://doi.org/10.1021/acsestwater.2c00209
10. Bonisławska, M., Nędzarek, A., Rybczyk, A., & Tański, A. (2023). The Influence of Anthropogenic Pollution on the Physicochemical Conditions of the Waters of the Lower Section of the Sąpólna River. Water, 16(1), 35. https://doi.org/10.3390/w16010035
11. Bozorg-Haddad, O., Delpasand, M., & Loáiciga, H. A. (2021). Water quality, hygiene, and health. In Economical, Political, and Social Issues in Water Resources (pp. 217–257). Elsevier. https://doi.org/10.1016/B978-0-323-90567-1.00008-5
12. Breiman, L. (2001). Random Forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324
13. Calvin, K., Dasgupta, D., Krinner, G., Mukherji, A., Thorne, P. W., Trisos, C., Romero, J., Aldunce, P., Barrett, K., Blanco, G., Cheung, W. W. L., Connors, S., Denton, F., Diongue-Niang, A., Dodman, D., Garschagen, M., Geden, O., Hayward, B., Jones, C., … Péan, C. (2023). IPCC, 2023: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental

Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland. (First). Intergovernmental Panel on Climate Change (IPCC). https://doi.org/10.59327/IPCC/AR6-9789291691647
14. Camara, M., Jamil, N. R., Abdullah, A. F. B., Hashim, R. B., & Aliyu, A. G. (2020). Economic and efficiency based optimisation of water quality monitoring network for land use impact assessment. Science of The Total Environment, 737, 139800. https://doi.org/10.1016/j.scitotenv.2020.139800
15. Cao, B., Gao, B., Liu, X., Wang, M., Yang, Z., & Yue, Q. (2011). The impact of pH on floc structure characteristic of polyferric chloride in a low DOC and high alkalinity surface water treatment. Water Research, 45(18), 6181–6188. https://doi.org/10.1016/j.watres.2011.09.019
16. Chen, S., Huang, J., Wang, P., Tang, X., & Zhang, Z. (2024). A coupled model to improve river water quality prediction towards addressing non-stationarity and data limitation. Water Research, 248, 120895. https://doi.org/10.1016/j.watres.2023.120895
17. Chen, W., Zhou, Z., He, J., Tao, H., & Liu, Z. (2017). Effect of typhoon with extreme precipitation on mountain reservoir drinking water treatment: A case study in Ningbo, China. Chinese Journal of Population Resources and Environment, 15(2), 103–108. https://doi.org/10.1080/10042857.2017.1319171
18. Chhipi-Shrestha, G., Mian, H. R., Mohammadiun, S., Rodriguez, M., Hewage, K., & Sadiq, R. (2023). Digital water: Artificial intelligence and soft computing applications for drinking water quality assessment. Clean Technologies and Environmental Policy, 25(5), 1409–1438. https://doi.org/10.1007/s10098-023-02477-4
19. Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). STL: A seasonal-trend decomposition procedure based on loess. Journal of Official Statistics, 6(1), 3-73.
20. Cleveland, W. S., & Devlin, S. J. (1988). Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting. Journal of the American Statistical Association, 83(403), 596–610. https://doi-org.ezproxy2.library.colostate.edu/10.1080/01621459.1988.10478639
21. Cody, T., Adams, S., & Beling, P. (2020). Motivating a Systems Theory of AI. INSIGHT, 23(1), 37–40. https://doi.org/10.1002/inst.12283
22. Czyczula Rudjord, Z., Reid, M. J., Schwermer, C. U., & Lin, Y. (2022). Laboratory Development of an AI System for the Real-Time Monitoring of Water Quality and Detection of Anomalies Arising from Chemical Contamination. WATER, 14(16), 2588. https://doi.org/10.3390/w14162588
23. Dallison, R. J. H., Williams, A. P., Harris, I. M., & Patil, S. D. (2022). Modelling the impact of future climate change on streamflow and water quality in Wales, UK. Hydrological Sciences Journal, 67(6), 939–962. https://doi.org/10.1080/02626667.2022.2044045
24. Denver Water & Coalition for the Upper South Platte. (2015). Source water protection plan for the Upper South Platte River (CO0116001). Denver Water.
25. DeMont, I., Anderson, L. E., Bennett, J. L., Sfynia, C., Bjorndahl, P., Jarvis, P., Stoddart, A. K., & Gagnon, G. A. (2024). Monitoring natural organic matter in drinking water treatment with photoelectrochemical oxygen demand. AWWA Water Science, 6(3), e1378. https://doi.org/10.1002/aws2.1378

26. Dickey, D. A., & Fuller, W. A. (1981). Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root. Econometrica, 49(4), 1057–1072. https://doi.org/10.2307/1912517
27. Wang, Z., & Yang, Y. (2024). Stationarity of high- and low-flows under climate change and human interventions across global catchments. Earth and Space Science, 11, e2023EA003456. https://doi.org/10.1029/2023EA003456
28. Eriksen, S., & Coelho, D. (2017). South Platte Watershed (Denver, Colorado). U.S. Environmental Protection Agency & U.S. Forest Service.
29. Fang, C., Yang, W., Lu, N., Xiao, R., Du, Z., Wang, Q., & Chu, W. (2023). Alkaline chlorination of drinking water: A trade-off between genotoxicity control and trihalomethane formation. Water Research, 246, 120692. https://doi.org/10.1016/j.watres.2023.120692
30. Farzana, S. Z., Paudyal, D. R., Chadalavada, S., & Alam, M. J. (2024). Decision Support Framework for Water Quality Management in Reservoirs Integrating Artificial Intelligence and Statistical Approaches. Water, 16(20), 2944. https://doi.org/10.3390/w16202944
31. Gauthier, V., Barbeau, B., Tremblay, G., Millette, R., & Bernier, A.-M. (2003). Impact of raw water turbidity fluctuations on drinking water quality in a distribution system. Journal of Environmental Engineering & Science, 2(4), 281–291. https://doi.org/10.1139/s03-026
32. Górniak, A. (2020). Total Organic Carbon in the Water of Polish Dam Reservoirs. In E. Korzeniewska & M. Harnisz (Eds.), Polish River Basins and Lakes – Part I (Vol. 86, pp. 189–207). Springer International Publishing. https://doi.org/10.1007/978-3-030-12123-5_10
33. Guinea, A. U., García, J., Cabero, J., Paunero, S., Bartolomé, M., Hernando, L. M., Maeso, P., & Benito, V. (2024). Adsorption and desorption processes of trihalomethanes on different granulated activated carbons in a full-scale advanced water treatment plant. Water Supply, 24(1), 1–10. https://doi.org/10.2166/ws.2023.324
34. Guo, D., Lintern, A., Webb, J. A., Ryu, D., Liu, S., Bende-Michl, U., Leahy, P., Wilson, P., & Western, A. W. (2019). Key Factors Affecting Temporal Variability in Stream Water Quality. Water Resources Research, 55(1), 112–129. https://doi.org/10.1029/2018WR023370
35. Guo, H., Song, Y., Tang, H., & Zhao, J. (2022). An ensemble deep neural network approach for predicting TOC concentration in lakes along the middle-lower reaches of Yangtze River. Journal of Intelligent & Fuzzy Systems, 42(3), 1455–1482. https://doi.org/10.3233/JIFS-210708
36. Haider, S., Rashid, M., Tariq, M. A. U. R., & Nadeem, A. (2024). The role of artificial intelligence (AI) and Chatgpt in water resources, including its potential benefits and associated challenges. Discover Water, 4(1), 113. https://doi.org/10.1007/s43832-024-00173-y
37. Heidari, H., Arabi, M., Warziniack, T., & Sharvelle, S. (2021). Effects of Urban Development Patterns on Municipal Water Shortage. Frontiers in Water, 3, 694817. https://doi.org/10.3389/frwa.2021.694817
38. Hirsch, R. M. (2011). A Perspective on Nonstationarity and Water Management1: A Perspective on Nonstationarity and Water Management. JAWRA Journal of the

American Water Resources Association, 47(3), 436–446. https://doi.org/10.1111/j.1752-1688.2011.00539.x
39. Singh, A., Sharma, P. J., & Teegavarapu, R. S. V. (2024). Understanding non-stationarity patterns in basin-scale hydroclimatic extremes. International Journal of Climatology, 44(11), 3867–3887. https://doi.org/10.1002/joc.8557
40. Jones, K. W., Cannon, J. B., Saavedra, F. A., Kampf, S. K., Addington, R. N., Cheng, A. S., MacDonald, L. H., Wilson, C., & Wolk, B. (2017). Return on investment from fuel treatments to reduce severe wildfire and erosion in a watershed investment program in Colorado. Journal of Environmental Management, 198, 66–77. https://doi.org/10.1016/j.jenvman.2017.05.023
41. J. Raseman, W., R. Kasprzyk, J., L. Rosario-Ortiz, F., R. Stewart, J., & Livneh, B. (2017). Emerging investigators series: A critical review of decision support systems for water treatment: making the case for incorporating climate change and climate extremes. Environmental Science: Water Research & Technology, 3(1), 18–36. https://doi.org/10.1039/C6EW00121A
42. Khan, S. J., Deere, D., Leusch, F. D. L., Humpage, A., Jenkins, M., & Cunliffe, D. (2015). Extreme weather events: Should drinking water quality management systems adapt to changing risk profiles? Water Research, 85, 124–136. https://doi.org/10.1016/j.watres.2015.08.018
43. Koley, S., Rao, K. B., Khwairakpam, M., & Kalamdhad, A. S. (2024). Identification and assessment of Critical parameters affecting drinking water quality: A case study of water treatment plants of India. Groundwater for Sustainable Development, 26, 101221. https://doi.org/10.1016/j.gsd.2024.101221
44. Konapala, G., Mishra, A. K., Wada, Y., & Mann, M. E. (2020). Climate change will affect global water availability through compounding changes in seasonal precipitation and evaporation. Nature Communications, 11(1), 3044. https://doi.org/10.1038/s41467-020-16757-w
45. Krbavčić, M., Sušanj Čule, I., Zorko, S., & Volf, G. (2023). Prediction models for manganese, iron and ammonium in raw water for a drinking water treatment plant Butoniga (croatia). Engineering Review, 43(3), 68–80. https://doi.org/10.30765/er.2232
46. Kumari, M., & Gupta, S. K. (2022). Cumulative human health risk analysis of trihalomethanes exposure in drinking water systems. Journal of Environmental Management, 321, 115949. https://doi.org/10.1016/j.jenvman.2022.115949
47. Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. Journal of Econometrics, 54(1–3), 159–178. https://doi.org/10.1016/0304-4076(92)90104-Y
48. Lap, B. Q., Phan, T.-T.-H., Nguyen, H. D., Quang, L. X., Hang, P. T., Phi, N. Q., Hoang, V. T., Linh, P. G., & Hang, B. T. T. (2023). Predicting Water Quality Index (WQI) by feature selection and machine learning: A case study of An Kim Hai irrigation system. Ecological Informatics, 74, 101991. https://doi.org/10.1016/j.ecoinf.2023.101991
49. Lehmann, N., Lantuit, H., Böttcher, M. E., Hartmann, J., Eulenburg, A., & Thomas, H. (2023). Alkalinity generation from carbonate weathering in a silicate-dominated headwater catchment at Iskorasfjellet, northern Norway. Biogeosciences, 20(16), 3459–3479. https://doi.org/10.5194/bg-20-3459-2023

50. Lei, X., Gao, L., Ma, M., Wei, J., Xu, L., Wang, L., & Lin, H. (2021). Does non-stationarity of extreme precipitation exist in the Poyang Lake Basin of China? Journal of Hydrology: Regional Studies, 37, 100920. https://doi.org/10.1016/j.ejrh.2021.100920
51. Lek, S., & Park, Y. S. (2008). Artificial Neural Networks. In Encyclopedia of Ecology, Five-Volume Set (Vol. 1-5, pp. 237-245). Elsevier. https://doi.org/10.1016/B978-008045405-4.00173-7
52. Lyle, Z. J. J., VanBriesen, J. M. M., & Samaras, C. (2023). Drinking Water Utility-Level Understanding of Climate Change Effects to System Reliability. ACS ES&T WATER, 3(8), 2395–2406. https://doi.org/10.1021/acsestwater.3c00091
53. McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysics, 5(4), 115–133. https://doi.org/10.1007/BF02478259
54. Melo, L. V., de Oliveira, M. D., Libânio, M., & Oliveira, S. C. (2016). Applicability of statistical tools for evaluation of water treatment plants. Desalination and Water Treatment, 57(30), 14024–14033. https://doi.org/10.1080/19443994.2015.1072586
55. Mensah-Akutteh, H., Buamah, R., Wiafe, S., & Nyarko, K. B. (2022). Raw water quality variations and its effect on the water treatment processes. Cogent Engineering, 9(1). https://doi.org/10.1080/23311916.2022.2122152
56. Menya, E., Olupot, P. W., Storz, H., Lubwama, M., & Kiros, Y. (2018). Production and performance of activated carbon from rice husks for removal of natural organic matter from water: A review. Chemical Engineering Research and Design, 129, 271–296. https://doi.org/10.1016/j.cherd.2017.11.008
57. Millar, G. J., Couperthwaite, S. J., & Moodliar, C. D. (2016). Strategies for the management and treatment of coal seam gas associated water. Renewable and Sustainable Energy Reviews, 57, 669–691. https://doi.org/10.1016/j.rser.2015.12.087
58. Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., & Stouffer, R. J. (2008). Stationarity Is Dead: Whither Water Management? Science, 319(5863), 573–574. https://doi.org/10.1126/science.1151915
59. Mott MacDonald, Hensyl, B. F., Borhani, S., Jacobs, Payab, A. H., Drexel University, Montalto, F., & Drexel University. (2024). Opportunities for Leveraging Existing Hydrologic and Hydraulic Models Developed for Water Quantity Management to Mitigate Flooding Due to Extreme Precipitation. Journal of Water Management Modeling. https://doi.org/10.14796/JWMM.C516
60. Murphy, J., & Sprague, L. (2019). Water-quality trends in US rivers: Exploring effects from streamflow trends and changes in watershed management. Science of The Total Environment, 656, 645–658. https://doi.org/10.1016/j.scitotenv.2018.11.255
61. Murphy, K. W., & Ellis, A. W. (2014). An assessment of the stationarity of climate and stream flow in watersheds of the Colorado River Basin. Journal of Hydrology, 509, 454–473. https://doi.org/10.1016/j.jhydrol.2013.11.056
62. Nabors, A., Barron, P., & Cochran, J. (2011). Using a Pilot Plant to demonstrate how raw water alkalinity can influence the treatment of drinking water. 319–328. https://doi.org/10.2495/WRM110271
63. Nemani, K. S., Peldszus, S., & Huck, P. M. (2023). Practical Framework for Evaluation and Improvement of Drinking Water Treatment Robustness in Preparation for Extreme-Weather-Related Adverse Water Quality Events. ACS ES&T Water, 3(5), 1305–1313. https://doi.org/10.1021/acsestwater.2c00627

64. Office of Cyber and Infrastructure Analysis (OCIA). (2015). Potential impacts of climate change on infrastructure in the Platte River Basin. National Protection and Programs Directorate, U.S. Department of Homeland Security.
65. Pesticides in surface water in agricultural and urban areas of the South Platte River basin, from Denver, Colorado, to North Platte, Nebraska, 1993-94. (1998). https://doi.org/10.3133/wri974230
66. Price, J. I., & Heberling, M. T. (2018). The Effects of Source Water Quality on Drinking Water Treatment Costs: A Review and Synthesis of Empirical Literature. Ecological Economics, 151, 195–209. https://doi.org/10.1016/j.ecolecon.2018.04.014
67. Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. Nature, 566(7743), 195–204. https://doi.org/10.1038/s41586-019-0912-1
68. U.S. Environmental Protection Agency. (2004). Report to Congress on impacts and control of combined sewer overflows and sanitary sewer overflows (EPA 833-R-04-001). Office of Water. Retrieved from http://www.epa.gov/npdes
69. Richards, C. E., Tzachor, A., Avin, S., & Fenner, R. (2023). Rewards, risks and responsible deployment of artificial intelligence in water systems. Nature Water, 1(5), 422–432. https://doi.org/10.1038/s44221-023-00069-6
70. Rizzo, L., Belgiorno, V., & Meriç, S. (2004). Organic THMs precursors removal from surface water with low TOC and high alkalinity by enhanced coagulation. Water Supply, 4(5–6), 103–111. https://doi.org/10.2166/ws.2004.0098
71. Rouse, W. B. (2020). AI as Systems Engineering Augmented Intelligence for Systems Engineers. INSIGHT, 23(1), 52–54. https://doi.org/10.1002/inst.12286
72. Rubio-Martin, A., Llario, F., Garcia-Prats, A., Macian-Sorribes, H., Macian, J., & Pulido-Velazquez, M. (2023). Climate services for water utilities: Lessons learnt from the case of the urban water supply to Valencia, Spain. Climate Services, 29, 100338. https://doi.org/10.1016/j.cliser.2022.100338
73. Salles, R., Belloze, K., Porto, F., Gonzalez, P. H., & Ogasawara, E. (2019). Nonstationary time series transformation methods: An experimental review. Knowledge-Based Systems, 164, 274–291. https://doi.org/10.1016/j.knosys.2018.10.041
74. Samson, C. C., Rajagopalan, B., & Summers, R. S. (2016). Modeling Source Water TOC Using Hydroclimate Variables and Local Polynomial Regression. Environmental Science & Technology, 50(8), 4413–4421. https://doi.org/10.1021/acs.est.6b00639
75. Serinaldi, F., Kilsby, C. G., & Lombardo, F. (2018). Untenable nonstationarity: An assessment of the fitness for purpose of trend tests in hydrology. Advances in Water Resources, 111, 132–155. https://doi.org/10.1016/j.advwatres.2017.10.015
76. Sheffield, J., Wood, E. F., & Roderick, M. L. (2012). Little change in global drought over the past 60 years. Nature, 491(7424), 435–438. https://doi.org/10.1038/nature11575
77. Shetty, A., & Goyal, A. (2022). Total organic carbon analysis in water – A review of current methods. Materials Today: Proceedings, 65, 3881–3886. https://doi.org/10.1016/j.matpr.2022.07.173
78. Shi, T., & Wu, J. (2021). Application of Artificial Intelligence in Water Conservancy Project Management. 2021 2nd International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE), 556–559. https://doi.org/10.1109/ICBASE53849.2021.00109

79. Shi, W., & Xia, J. (2017). Combined risk assessment of nonstationary monthly water quality based on Markov chain and time-varying copula. Water Science and Technology, 75(3), 693–704. https://doi.org/10.2166/wst.2016.553
80. Sillanpää, M., Matilainen, A., & Lahtinen, T. (2015). Characterization of NOM. In Natural Organic Matter in Water (pp. 17–53). Elsevier. https://doi.org/10.1016/B978-0-12-801503-2.00002-1
81. Skaland, R. G., Herrador, B. G., Hisdal, H., Hygen, H. O., Hyllestad, S., Lund, V., White, R., Wong, W. K., & Nygård, K. (2022). Impacts of climate change on drinking water quality in Norway. Journal of Water and Health, 20(3), 539–550. https://doi.org/10.2166/wh.2022.264
82. Slater, L. J., Anderson, B., Buechel, M., Dadson, S., Han, S., Harrigan, S., Kelder, T., Kowal, K., Lees, T., Matthews, T., Murphy, C., & Wilby, R. L. (2021). Nonstationary weather and water extremes: A review of methods for their detection, attribution, and management. Hydrology and Earth System Sciences, 25(7), 3897–3935. https://doi.org/10.5194/hess-25-3897-2021
83. Szomolányi, O., & Clement, A. (2023). Use of random forest for assessing the effect of water quality parameters on the biological status of surface waters. GEM - International Journal on Geomathematics, 14(1), 20. https://doi.org/10.1007/s13137-023-00229-6
84. Szpak, D., Barbara, Tchorzewska-Cieslak, & Pietrucha-Urbanik, K. (2020). Analysis of the turbidity of raw water in the context of water-supply safety. Desalination and Water Treatment, 186, 281–289. https://doi.org/10.5004/dwt.2020.25176
85. Terry, J., & Lindenschmidt, K.-E. (2023). Modelling Climate Change and Water Quality in the Canadian Prairies Using Loosely Coupled WASP and CE-QUAL-W2. WATER, 15(18), 3192. https://doi.org/10.3390/w15183192
86. Tibshirani, R. (2023). Lecture 2: Measures of dependence and stationarity. Introduction to Time Series, Fall 2023. Retrieved from https://www.stat.berkeley.edu/~ryantibs/timeseries-f23/lectures/dependence.pdf.
87. Towler, E., Rajagopalan, B., & Summers, R. S. (2009). Using Parametric and Nonparametric Methods to Model Total Organic Carbon, Alkalinity, and pH after Conventional Surface Water Treatment. Environmental Engineering Science, 26(8), 1299–1308. https://doi.org/10.1089/ees.2008.0341
88. Tu, J. (2009). Combined impact of climate and land use changes on streamflow and water quality in eastern Massachusetts, USA. Journal of Hydrology, 379(3–4), 268–283. https://doi.org/10.1016/j.jhydrol.2009.10.009
89. Ubah, J. I., Orakwe, L. C., Ogbu, K. N., Awu, J. I., Ahaneku, I. E., & Chukwuma, E. C. (2021). Forecasting water quality parameters using artificial neural network for irrigation purposes. Scientific Reports, 11(1), 24438. https://doi.org/10.1038/s41598-021-04062-5
90. Wang, F., Wang, Y., Zhang, K., Hu, M., Weng, Q., & Zhang, H. (2021). Spatial heterogeneity modeling of water quality based on random forest regression and model interpretation. Environmental Research, 202, 111660. https://doi.org/10.1016/j.envres.2021.111660
91. Wang, L., Liu, J., Zhao, Q., Wei, W., & Sun, Y. (2016). Comparative study of wastewater treatment and nutrient recycle via activated sludge, microalgae and combination systems. Bioresource Technology, 211, 1–5. https://doi.org/10.1016/j.biortech.2016.03.048

92. World Meteorological Organization (WMO). (2024). State of climate services: Five-year progress report (2019–2024). WMO. https://public.wmo.int/en/resources/library
93. Wright, B., Stanford, B. D., Reinert, A., Routt, J. C., Khan, S. J., & Debroux, J. F. (2014). Managing water quality impacts from drought on drinking water supplies. Journal of Water Supply: Research and Technology-Aqua, 63(3), 179–188. https://doi.org/10.2166/aqua.2013.123
94. Wu, W., Dandy, G. C., & Maier, H. R. (2014). Protocol for developing ANN models and its application to the assessment of the quality of the ANN model development process in drinking water quality modelling. Environmental Modelling & Software, 54, 108–127. https://doi.org/10.1016/j.envsoft.2013.12.016
95. Xiao, R., Yang, X., Fang, C., Zhang, R., & Chu, W. (2023). Total organic halogen (TOX) in drinking water: Occurrence, correlation analysis, and precursor removal during drinking water treatment. Science of The Total Environment, 905, 167445. https://doi.org/10.1016/j.scitotenv.2023.167445
96. Xu, J., Xu, Z., Kuang, J., Lin, C., Xiao, L., Huang, X., & Zhang, Y. (2021). An Alternative to Laboratory Testing: Random Forest-Based Water Quality Prediction Framework for Inland and Nearshore Water Bodies. Water, 13(22), 3262. https://doi.org/10.3390/w13223262
97. Yang, Y., Roderick, M. L., Yang, D., Wang, Z., Ruan, F., McVicar, T. R., Zhang, S., & Beck, H. E. (2021). Streamflow stationarity in a changing world. Environmental Research Letters, 16(6), 064096. https://doi.org/10.1088/1748-9326/ac08c1
98. Yee Wong, W., Khallel Ibrahim Al-Ani, A., Hasikin, K., Salwa Mohd Khairuddin, A., Abdul Razak, S., Farzana Hizaddin, H., Istajib Mokhtar, M., & Mokhzaini Azizan, M. (2022). Water Quality Index Using Modified Random Forest Technique: Assessing Novel Input Features. Computer Modeling in Engineering & Sciences, 132(3), 1011–1038. https://doi.org/10.32604/cmes.2022.019244
99. Yilmaz, A. G., Imteaz, M. A., Shanableh, A., Al-Ruzouq, R., Atabay, S., & Haddad, K. (2023). A Non-Stationarity Analysis of Annual Maximum Floods: A Case Study of Campaspe River Basin, Australia. Water, 15(20), 3683. https://doi.org/10.3390/w15203683
100. Yirenkyi-Fianko, A. B., Yanful, E. K., & Ottou, J. A. (2022). Seasonal variation of water quality parameters of surface water in mining areas. Management of Environmental Quality: An International Journal, 33(5), 1290–1304. https://doi.org/10.1108/MEQ-02-2022-0037
101. Younos, T., & Grady, C. A. (Eds.). (2014). Potable Water: Emerging Global Problems and Solutions (Vol. 30). Springer International Publishing. https://doi.org/10.1007/978-3-319-06563-2
102. Yuan, X., & O'Loughlin, F. (2024). Evaluating the Non-Stationarity, Seasonality and Temporal Risk to Water Resources in the Wei River Basin. Water, 16(17), 2513. https://doi.org/10.3390/w16172513
103. Yuan, X., Wang, Y., Ji, P., Wu, P., Sheffield, J., & Otkin, J. A. (2023). A global transition to flash droughts under climate change. Science, 380(6641), 187–191. https://doi.org/10.1126/science.abn6301
104. Zavareh, M., Maggioni, V., & Zhang, X. (2024). Assessing the Efficiency of a Random Forest Regression Model for Estimating Water Quality Indicators. Meteorology Hydrology and Water Management. https://doi.org/10.26491/mhwm/183734

105. Zhang, K., Achari, G., Sadiq, R., Langford, C. H., & Dore, M. H. I. (2012). An integrated performance assessment framework for water treatment plants. Water Research, 46(6), 1673–1683. https://doi.org/10.1016/j.watres.2011.12.006
106. Zhang, Q., Gu, X., Singh, V. P., Xiao, M., & Xu, C.-Y. (2014). Stationarity of annual flood peaks during 1951–2010 in the Pearl River basin, China. Journal of Hydrology, 519, 3263–3274. https://doi.org/10.1016/j.jhydrol.2014.10.028
107. Zhang, Q., & You, X. (2024). Recent Advances in Surface Water Quality Prediction Using Artificial Intelligence Models. Water Resources Management, 38(1), 235–250. https://doi.org/10.1007/s11269-023-03666-y
108. Zhang, Y., Yao, X., Wu, Q., Huang, Y., Zhou, Z., Yang, J., & Liu, X. (2021). Turbidity prediction of lake-type raw water using random forest model based on meteorological data: A case study of Tai lake, China. Journal of Environmental Management, 290, 112657. https://doi.org/10.1016/j.jenvman.2021.112657

LIST OF ABBREVIATIONS

AI – Artificial Intelligence
ML – Machine Learning
DWTP – Drinking Water Treatment Plant
RWQS – Raw Water Quality State
TOC – Total Organic Carbon
NOM – Natural Organic Matter
DOM – Dissolved Organic Matter
ADF – Augmented Dickey-Fuller Test
KPSS – Kwiatkowski-Phillips-Schmidt-Shin Test
STL – Seasonal-Trend Decomposition Using Loess
DBPs – Disinfection By-Products
NTU – Nephelometric Turbidity Units
THMs – Trihalomethanes
DOC – Dissolved Organic Carbon
TA – Total Alkalinity
ACC – Accuracy
DL – Deep Learning
MPN – Most Probable Number
NFSPR – North Fork of the South Platte River
LOESS – Locally Weighted Regression
EFA – Exploratory Factor Analysis
RF – Random Forest
MSE – Mean Squared Error
MAE – Mean Absolute Error
GridSearchCV – Grid Search with 5-Fold Cross-Validation
FNN – Feedforward Artificial Neural Network
MLP – Multilayer Perceptron