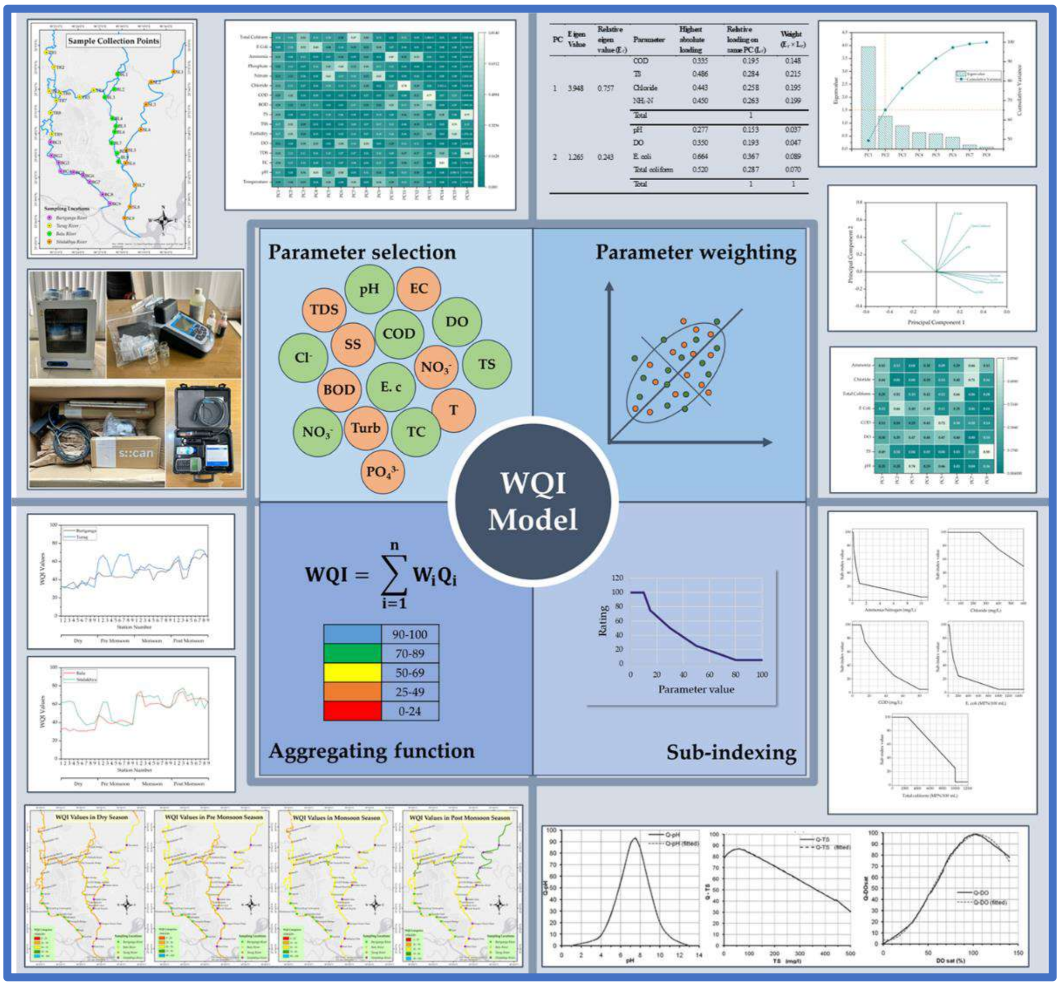
Government of the People’s Republic of Bangladesh
Ministry of Water Resources
Water Resources Planning Organization (WARPO)



Establishment of Water Quality Index (WQI) through Principal Component Analysis for the Dhaka-based Rivers

Final Report
June, 2025

Water Resources Planning Organization

Department of Chemical Engineering, BUET

Government of the People's Republic of Bangladesh
Ministry of Water Resources
Water Resources Planning Organization (WARPO)

Establishment of Water Quality Index (WQI) through Principal Component Analysis for the Dhaka-based Rivers

Final Report
June, 2025

Water Resources Planning Organization

Department of Chemical Engineering, BUET

Table of Content

List of Figures

List of Tables

Executive Summary

Rivers are essential lifelines for transportation, agriculture, fisheries, industry, and daily human activities. However, rapid urbanization, unregulated industrial expansion, and untreated wastewater discharge have severely deteriorated water quality, particularly in rivers surrounding Dhaka city. The Buriganga, Turag, Balu, and Shitalakhya rivers, which once sustained the economy and ecology of the capital Dhaka, are now heavily polluted due to industrial effluents, sewage dumping, and encroachments. Excessive organic loads, toxic heavy metals, and microbial contamination have pushed these rivers to critical levels, endangering aquatic ecosystems and posing severe risks to human health. Conventional water quality monitoring methods often involve assessing multiple parameters, making data interpretation complex and resource-intensive. To address this, the Water Resources Planning Organization (WARPO) and the Department of Chemical Engineering, BUET, initiated this study to develop a scientifically validated Water Quality Index (WQI) model using Principal Component Analysis (PCA). This project aims to establish a structured and efficient method to quantify water quality, identify key pollution drivers, and support data-driven policymaking. The model's application will enable continuous monitoring, improve pollution management strategies, and provide a foundation for regulatory interventions to restore the health of Dhaka's River systems.

The study focuses on analyzing primary water quality data collected from 36 strategically selected monitoring stations across four different seasons—dry, pre-monsoon, monsoon, and post-monsoon—to capture seasonal variations. The dataset includes measurements of temperature, pH, electrical conductivity (EC), total dissolved solids (TDS), total suspended solids (TSS), total solids (TS), dissolved oxygen (DO), chemical oxygen demand (COD), ammonia-nitrogen, chloride, phosphate, nitrate, turbidity, biological oxygen demand (BOD), and microbial contaminants (E. coli, total coliform). Principal Component Analysis (PCA) was applied to reduce dimensionality and identify the most critical water quality indicators. The first three principal components explained 74% of the total variance, supporting the selection of a reduced set of key parameters. Along with PCA and correlation analysis, eight key parameters were selected for WQI development. For sub-indexing, both quality rating curves and linear interpolation functions were used to convert raw parameter values into a standardized score.

The developed WQI model reveals significant seasonal variations in water quality across the four rivers. Among the four rivers, the Buriganga and Turag exhibited the poorest water quality, particularly during the dry season, when dilution capacity is at its lowest. The lowest WQI values were recorded at Fulpukuria Thread (Turag) with 29.78 and Chandir Ghat (Buriganga) with 30.50, categorizing them as "Bad" water quality. The Shitalakhya and Balu rivers have relatively better water quality. The Balu River records its lowest WQI of 30.52 (Talia) during the dry season, while its highest WQI reaches 74.29 (Ulukhola) in the monsoon season, classifying it as "Good" water quality. Similarly, the Shitalakhya River has its lowest WQI of 36.49 (Horipur) in the pre-monsoon season and its highest WQI of 77.96 (Beldi Bazar) in the post-monsoon season, also indicating "Good" water quality. Both rivers generally exhibit improved water quality during the monsoon and post-monsoon seasons due to increased flow and dilution. Sobol sensitivity analysis identified TS, COD, and chloride are the most sensitive parameters, meaning they have the highest influence on the WQI score. DO and ammonia-nitrogen showed significant interaction effects, meaning changes in these parameters could indirectly impact overall water quality.

To prevent further degradation and improve river water quality, strict enforcement of wastewater regulations and enhanced environmental monitoring are essential. Real-time monitoring, public awareness, and stakeholder engagement will support effective pollution control. Future research should include long-term data collection and assess emerging pollutants for a more comprehensive understanding.

This study presents a scientifically sound, cost-effective, and adaptable WQI model for assessing water quality in the major rivers of Dhaka city. The use of PCA and correlation analysis-driven parameter selection ensures a more efficient and statistically valid assessment of water quality trends. The findings emphasize the urgent need for regulatory actions, pollution control measures, and public awareness campaigns to prevent further environmental deterioration. With continued monitoring and policy support, this model can serve as a valuable tool for decision-makers, enabling more effective water quality management and restoration efforts in Bangladesh.

Chapter 1: Introduction

1.1 Background

River water is one of the most useful resources to people because of its multipurpose uses including domestic, industrial and agricultural in addition to transportation, tourism and recreational purposes. Dhaka division is one of the eight administrative divisions of Bangladesh which includes the Dhaka city, the capital of Bangladesh. The current population of the city is about 20 million, which turns Dhaka city into one of the most densely populated megacities in the world. Considering the importance of rivers to humans, Dhaka city has grown on the bank of Buriganga from the era of the Mughal empire and is surrounded by four rivers (Figure 1.1). There are Balu and Sitalakhya on the eastern side, Turag and Buriganga on the western side. These rivers receive water from the Jamuna (Brahmaputra River) in the wet season, and in the dry season upper reaches of these rivers are slowly replenished by the release of groundwater into the rivers. The lower reaches of the rivers are also influenced by the tidal variations travelling upstream from the Bay of Bengal. In the monsoon season, the river levels reach around 6.5m MSL (mean sea level) and drop to about 2.5m MSL in the dry season (BWDB, 2023). However, the unplanned development of the city and the establishment of different industries on the banks of the rivers over time resulted in a decrease in the water quality of the rivers.

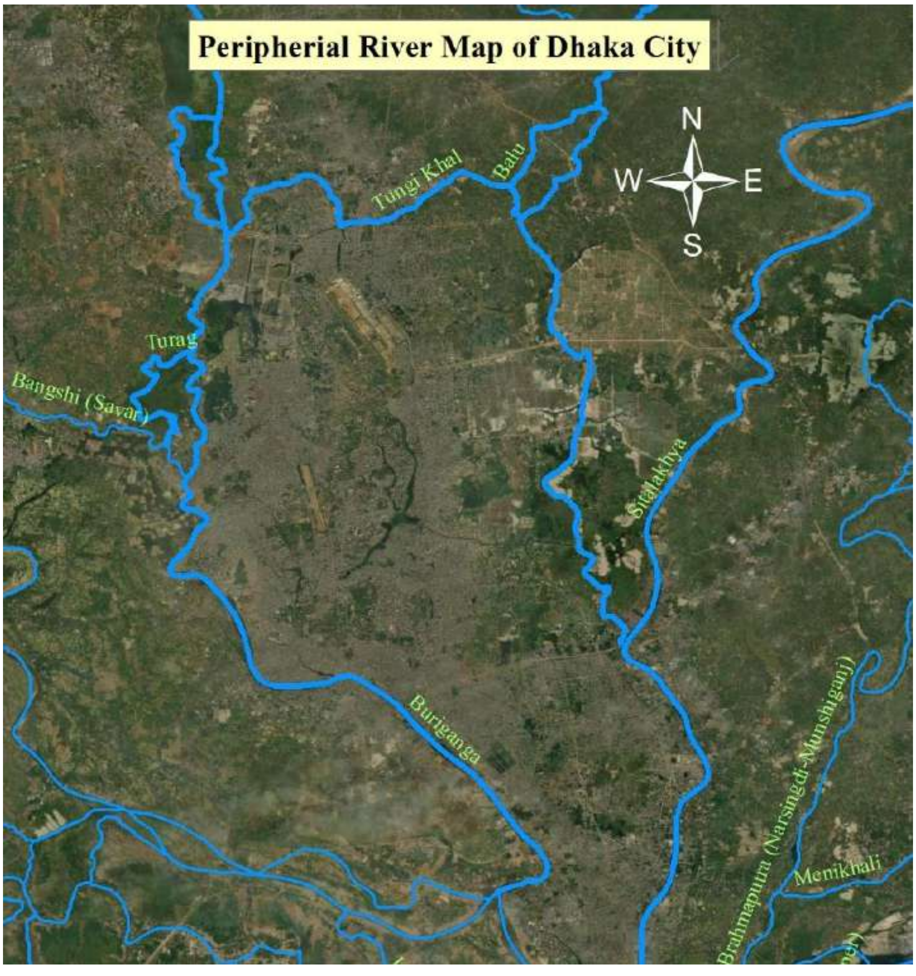


Figure 1.1: Peripheral Rivers of Dhaka City

Figure 1.2: The pollution scenario of Buriganga, Turag, Balu and Shitalakhya rivers

There are a few studies available on the Dhaka-based River water quality assessment, which were collected and reviewed. Mir Mostafa Kamal et al. reported the alarmingly low level of DO in the Buriganga river back in the 2000s (Kamal et al., 1999). Many other researchers studied water quality parameters e.g., dissolved oxygen (DO), biochemical oxygen demand (BOD), chemical oxygen demand (COD), pH, turbidity, conductivity, total dissolved solids (TDS), nitrate and phosphate in Buriganga river and those studies found the DO, BOD, COD, TDS, turbidity, nitrate and phosphate are at an alarming level and a discussion on the possible sources of the pollution are presented in some of the papers (Akbor et al., 2017; Fatema et al., 2018). A study indicates the pollution sources of Buriganga rivers are both point and non-point, making the river water quality highly unhealthy from the perspective of aquatic ecosystems (Salman et al., 2018). Another study on the water quality assessment of the Buriganga river conducted by Md. Ashiqur Rahman et al. discussed the effect of different water quality parameters and concluded most of the parameters are not acceptable according to the allowable limit during both dry and wet seasons (M. A. Rahman & Al Bakri, 2010). Shaikh et al. studied 10 parameters for 10 different sampling stations of the Buriganga river to evaluate the water quality and found unsatisfactory results for each station (Sayed Ahammed et al., 2016). Also, there have been researches done on other peripheral rivers of Dhaka city regarding water quality assessment (Naushad Alam et al., n.d.; Nazma Sultana et al., 2019; A. Rahman et al., 2012).

One of the most important objectives of water quality assessment studies is assessing the seasonal variation which has been addressed in most of the previous works. Several government and international agencies reported four seasons in Bangladesh. Dry/Winter, Pre-Monsoon, Monsoon and Post-Monsoon seasons. Table 1.1 indicates the above-mentioned four seasons in different literature from government and international agencies.

Table 1.1: Seasonal variation in Bangladesh

Season Name	Duration	Month Span	Agency Names	Reference
Dry/Winter	3 months	December-February	WARPO	(Ministry of Water Resource, 2001)
Pre-Monsoon	3 months	March-May	BMD	(Khatun et al., n.d.)
Monsoon	4 months	June-September	FAO	(*Country Profile-Bangladesh*, 2014)
Post-Monsoon	2 months	October-November	World Bank	(Country Profile Bangladesh COUNTRY OVERVIEW, 2011)

But in contrast to these reports, several scholars have studied the water quality of rivers in Bangladesh mentioning different timespan for different seasons. For example, a study on water quality assessment of Sitalkhya river was conducted mentioning the seasonal variation of two seasons, dry and wet (Naushad Alam et al., n.d.). However, there was no specific month or duration mentioned about the seasonal span. Other studies were conducted in the peripheral rivers of Dhaka city, especially Buriganga and Turag, and most of them mentioned dry and wet seasons (Fatema et al., 2018; Kamal et al., 1999; A. Rahman et al., 2012; S. Rahman & Hossain, 2008). Some of them specified the seasonal span, for instance, Fatema et al. mentioned two seasons, dry (November – January) and wet (June – August) (Fatema et al., 2018) from which existence of a total of four seasons in Bangladesh can be inferred and these two seasons can be pre-monsoon and post-monsoon mentioned in the *Development of an Assessment System to Evaluate the Ecological Status of Rivers in the Hindu Kush-Himalayan Region* (ASSESS-HKH) project (Moog & Scientific Conference Rivers in the Hindu Kush-Himalaya - Ecology & Environmental Assessment 2008 Kāṭhamāṇḍū; Dhulikhel, n.d.). In contrast, Rahman et. al. mentioned in their study about seasons covering 12 months of the year – dry season (December to May) and the wet season (June to November) (A. Rahman et al., 2012). Another Study based on the Buriganga river mentioned three conventional seasons – summer, winter and autumn and specified the months (Sayed Ahammed et al., 2016). Since there is no specified seasonal span, these studies differed from each other, but all these studies concluded a generalized result that water quality highly depends on seasonal variation. In this current work, we are aiming to establish a seasonal variation in water quality in accordance with the national and international reports regarding the seasons.

Table 1.2 represents a summary of the parameters used in different studies along with their findings.

Table 1.2: Parameters used in previous studies and their findings

Study Name	Parameters Studied	No. of Sampling Stations and Seasons	Approach and Major Findings	References
Assessment of Pollution of the River Buriganga, Bangladesh, Using a Water Quality Model	DO, BOD_5, COD, NH_3-N, NH_4-N, E. Coli, Total Coliform, TSS, NO_3-N, Orth.PO_4-P, Cr	Sampling was done in 1 Station during 1 Season (Dry)	Low DO and higher values of other parameters were found with an occasional rise to DO and low concentration of other parameters due to local rainfall and during the spring time period.	(Kamal et al., 1999)
Investigation of Water Quality Parameters at Different Points in the Buriganga River, Bangladesh	EC, TDS, Salinity, pH, DO, BOD, COD, Hardness, Acidity, Free CO_2, Alkalinity, Turbidity	Sampling was done in 8 Station during 1 Season (Dry)	High BOD, COD, EC, TDS, Salinity, Alkalinity and Turbidity with low DO indicating very critical water condition of river Buriganga.	(Akbor et al., 2017)
Water Quality Assessment of the River Buriganga, Bangladesh	Temperature, DO, pH, Conductivity, Phosphate, Mn, Fe, Pb, Cd	Sampling was done in 3 Station during 2 Season (Wet and Dry)	Significant temperature difference depends on the months, not on the sampling stations. Differences in pH, DO, Phosphate etc. are significant between dry and wet seasons.	(Fatema et al., 2018)
Water Quality Assessment of the Buriganga River, Dhaka, Bangladesh	pH, Temperature, Alkalinity, Color, Turbidity, Chloride, Iron, BOD	Sampling was done in 4 Station	All the water quality parameters measured fall below the standard and acceptable reason. Parameter values change from Hazaribag to Postogola.	(Salman et al., 2018)
A Study on Selected Water Quality Parameters along the River Buriganga, Bangladesh	Temperature, pH, EC, DO, BOD, COD, PO_4-P, NH_3-N, Pb, Cr	Sampling was done in 5 Station during 2 Season (Wet and Dry)	The observed DO was lower than the standard level for all sampling stations. Low DO could be linked to higher turbidity.	(M. A. Rahman & Al Bakri, 2010)

Study Name	Parameters Studied	No. of Sampling Stations and Seasons	Approach and Major Findings	References
Water Quality Assessment of Balu River, Dhaka Bangladesh	pH, EC, Fe, Cd, Pb, Cr	Sampling was done in 1 Station during 1 Season (Dry)	A significant number of heavy metals was found in the samples and the overall quality of the river is found to be threatening.	(Nazma Sultana et al., 2019)
Study of the seasonal variations in Turag river water quality parameters	pH, EC, Salinity, TDS, DO, BOD, COD, Free CO2, Hardness, As, Cd, Pb, Zn, Hg	Sampling was done in 3 Station during 2 Season (Wet and Dry)	This study was conducted to compare the seasonal effect on water quality parameters in dry and wet seasons. Higher BOD and COD were found during the dry season. TDS, EC and Hardness also increased during the dry season which fell below the standard limit.	(A. Rahman et al., 2012)
Risk and Water Quality Assessment overview of River Sitalakhya in Bangladesh	DO, pH, EC, TDS		Studied the parameter trend over the years in the Shitalakhya river. DO tends to decrease and TDS and EC are increasing each year.	(Naushad Alam et al., n.d.)
An Investigation Into The Water Quality Of Buriganga - A River Running Through Dhaka	DO, BOD, COD, pH, Temperature, Conductivity, Turbidity, TDS, Nitrate, Phosphate	Sampling was done in 10 Station during 3 Season (Summer, Winter and Autumn)	This study measured water quality parameters of selected sites and compared them with other literature values. Based on the parameter values probable point sources of pollution were indicated and during the dry season, most of the parameters deteriorated than the wet season.	(Sayed Ahammed et al., 2016)

Paul Whitehead et al. (2018) and his team performed a baseline survey of water chemistry and total coliforms of the Turag-Tongi-Balu river system and showed DO close to zero in the dry season, high organic loading together with extreme levels of Ammonium-N and total coliform in the water (Whitehead et al., 2018). However, most of the studies are scattered, non-comprehensive, dealing with several parameters and limited sites of the river. All these mentioned surface water quality assessment studies focused on a singular evaluation of parameters individually, no study was conducted to combine the important parameters and represent them as a single entity to evaluate water quality. These studies emphasized more insights, causes and consequences of river water pollution and thus, these studies could not reach the policymakers. In order to introduce long-term and effective river water management, focused and structured research is necessary to understand the actual condition of the study area (baseline). In order to recognize the Dhaka-based River water quality, a good source of data on the water quality is indispensable. Although the regular measurement of water quality parameters is a laborious job because of spatial and temporal variations of water environment quality along the river, it will provide a representative and reliable estimation of surface water quality. The long-term monitoring of many profiles with different reach would generate a large and complex database, which needs a good approach to interpretation (Kabir et al., 2022). To identify and mitigate the pollution source, the correct interpretation of the collected data is crucial. Moreover, random site selection of study areas is also an issue in water quality assessment. The geographical information system (GIS) is very helpful in evaluating the spatial distribution of water quality parameters over the study area (Oseke et al., 2021). The collected data can be analyzed by the application of different multivariate statistical techniques, such as Principal Component Analysis (PCA), Cluster Analysis (CA), and multiple linear regression. These techniques assist in the interpretation of complex data matrices to better understand the water quality and ecological status of the studied systems, allow the identification of possible factors that influence water environment systems and offer a valuable tool for the reliable management of water resources. Assessment of river water quality or optimization of the monitoring procedure of river water is linked to clustering of sampling locations, river water quality parameters, identification of possible sources of pollution, or modeling the contribution of the identified sources to the formation of the total concentration of the monitored chemical tracers.

Most of the time, the research on water quality assessment fails due to the thousands of output information because water quality regulations and monitoring programs generate a vast number of multidimensional water quality parameter sets, pollution characteristics, and numerical data about various water sources that are understandable only to the scientists (Oketola et al., 2013). This information, however, should be beneficial to the authorities of the water sector in decision-making to restore the water quality. As a result, a technique known as the Water Quality Index (WQI) has been developed to solve this issue. The Water Quality Index is a quantitative representation that is used to determine the ecological health of a body of water. The purpose of the Water Quality Index is to categorize waters according to their physical, chemical, and biological attributes, thereby establishing their potential uses and controlling their allocation decisions (Pesce & Wunderlin, 2000). By providing a single dimensionless value, the WQI helps to reduce the multivariate nature of the data that describes the quality of water (Tharmar et al., 2022) and specify the impact of seasonal variation in water quality. Because of its generalized structure, it has been a popular choice since the 1960s when it was first introduced (M. G. Uddin et al., 2021). More than 35 WQI models have been developed or introduced by different countries or agencies for the assessment/evaluation of surface water quality till now (Abbasi & Abbasi, 2012c; Dadolahi-Sohrab et al., 2012; Kannel et al., 2007; Stoner et al., n.d.).

Thus, proper site selection, proper data collection, correct analysis of the collected data, and indicative results all should be a part of a river water quality assessment project. Otherwise, the research would be limited to the researcher only. In this research, a comprehensive study will be carried out to understand four main river water characteristics using ArcGIS, numerical Principal Component Analysis (PCA) and WQI analysis.

1.2 Research Motivation

The water quality of rivers, a vital component of urban ecosystems, plays a pivotal role in sustaining both environmental health and human well-being. Despite the undeniable importance of this subject, the current state of research on water quality in Dhaka-based rivers reveals several critical gaps that necessitate immediate attention and comprehensive investigation.

1) **Lack of Sufficient Studies:** One of the primary motivations for this research stems from the apparent lack of sufficient studies addressing the water quality of Dhaka-based rivers. Existing literature reveals a paucity of in-depth examinations into the multifaceted factors influencing water quality, highlighting the urgent need for a more thorough investigation.
2) **Most of the Studies are Non-comprehensive:** Existing studies on water quality in the region often fall short in providing a comprehensive analysis. Many have focused on a limited set of parameters, neglecting the intricate interplay of various factors that collectively define water quality. This research aims to address this limitation by adopting a more holistic approach, considering a broader spectrum of influential variables.
3) **Dealt with a Vast Number of Parameters:** Contrary to the non-comprehensive nature of some studies, others have attempted to tackle water quality by incorporating an extensive array of parameters. However, this approach, while well-intentioned, often leads to information overload and lacks a systematic method to distill the most crucial factors. Our research seeks to streamline this process through the application of Principal Component Analysis (PCA), providing a more nuanced understanding of the interrelationships among diverse parameters.
4) **No Recommendations for Policy Makers:** Despite numerous studies on water quality, a notable gap exists in translating research findings into actionable recommendations for policymakers. Our research aspires not only to identify areas of concern but also to offer concrete suggestions and guidelines for policymakers to formulate effective strategies for water quality management in the Dhaka region.
5) **No Innovative Approach:** The absence of an innovative approach in current studies underscores the need for novel methodologies that can offer fresh perspectives and insights. By integrating Principal Component Analysis into the establishment of a Water Quality Index (WQI), our research introduces an innovative framework, fostering a more sophisticated understanding of the complex dynamics influencing river water quality.
6) **Development of a WQI for River Water:** Finally, our research is motivated by the aspiration to develop a Water Quality Index (WQI) for Dhaka-based rivers. Inspired by the success of the Air Quality Index (AQI), which has become a widely accepted and accessible

tool, a similar metric for water quality is crucial for raising public awareness, aiding policymakers, and facilitating informed decision-making in the context of environmental management.

In summary, the identified gaps and limitations in existing research on water quality in Dhaka-based rivers underscore the significance of this study. By addressing these shortcomings and introducing an innovative approach, our research aims to contribute substantially to the understanding and management of water quality, with the ultimate goal of establishing a practical Water Quality Index for the region. Figure 1.3 points out the major motivation behind this research study.

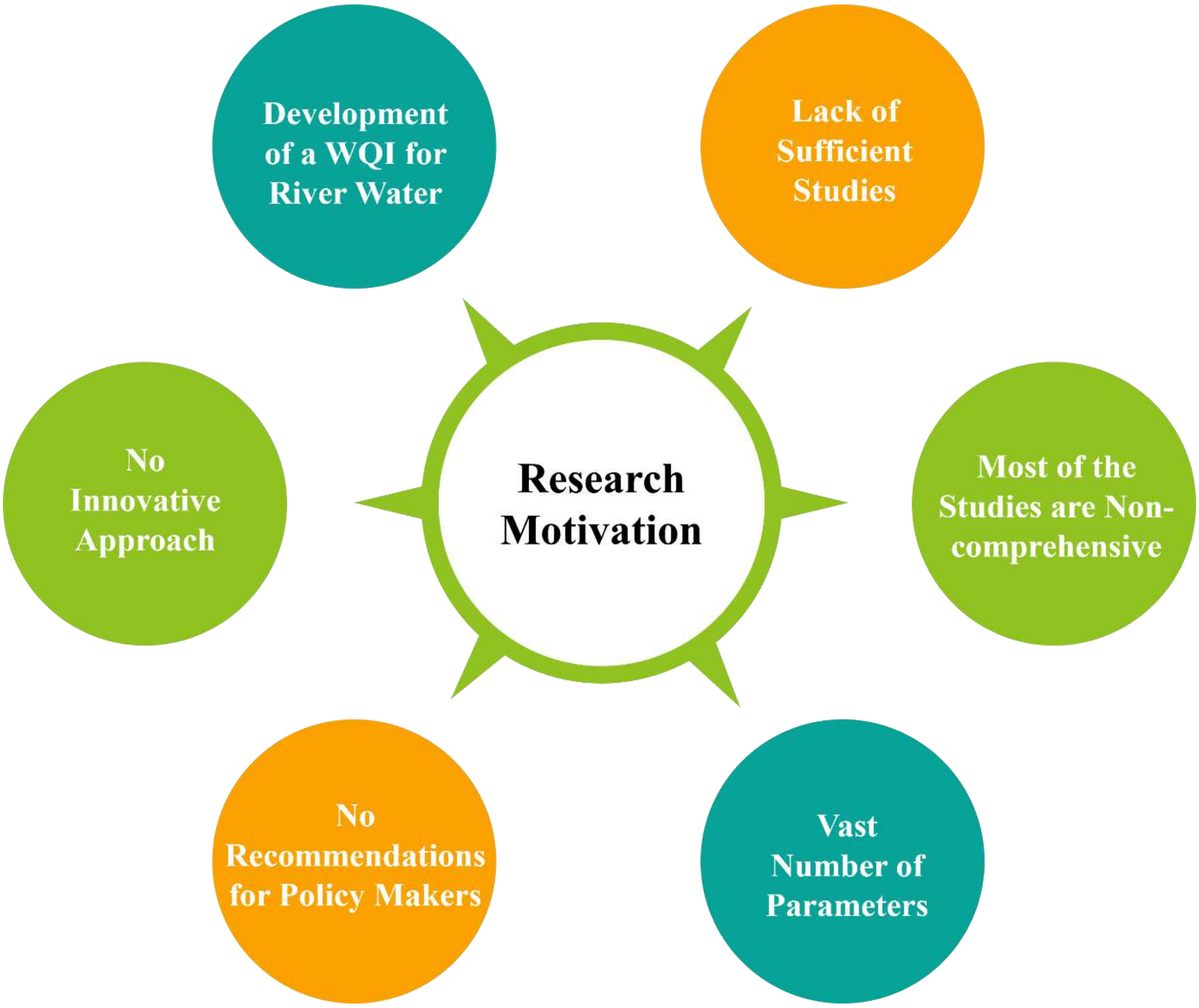


Figure 1.3: Motivation of the research project

1.3 Research Objectives

The main objective of the research is to establish a comprehensive method to determine the Water Quality Index (WQI) for the Dhaka-based rivers. On a pilot basis, the research would concentrate on the selected reaches of the peripheral rivers of Dhaka city - Buriganga, Turag, Balu and Shitalakshya rivers.

The following specific objectives have been set to achieve the goal of this research work:

- To understand the baseline water quality parameters of the Buriganga, Turag, Balu and Shitalakshya rivers.
- To identify the most critical water quality parameters for the selected rivers using principal component analysis (PCA).
- To establish a method, consistent with the internationally adopted approach, to calculate the water quality index (WQI) for the selected rivers.
- To categorize and classify the selected rivers considering seasonal variation and based on their respective WQI values.

1.4 Scope of Work

- Collect secondary water quality data of 4 rivers from National Water Resources Database (NWRD) and other secondary sources
- Select potential monitoring sites for sample collection for each river through field survey
- Collect data for water quality parameters (pH, TDS, TSS, EC, BOD, COD, DO etc.) and water samples following standard protocols
- Perform laboratory tests for other important water quality parameters (amount of different ions like phosphate, nitrate, chloride, and ammonia, amount of E. coli and total coliform organisms etc.)
- Analyze the parameters by Principal Component Analysis (PCA) to obtain the critical water quality parameters
- Derive an appropriate method to calculate water quality index (WQI) for the selected rivers

1.5 Research Team Composition

This is a collaborative research project with the participation of the Department of Chemical Engineering (DChE), BUET and WARPO. BUET team consists of Principal Investigator, 2 Research Assistants and 1 professional with relevant background. WARPO team consists of 4 professionals with relevant backgrounds. The Research Team composition is shown in Table 1.3.

Table 1.3: Research team composition

Name	Designation, Institution and Position in the Research Team
Prof. Dr. Md. Shahinoor Islam	Professor, DChE, BUET & Principal Investigator
Kazi Saidur Rahman	Senior Scientific Officer, WARPO & Principal Research Coordinator
Md. Al-Amin Kabir Bhuiyan	Senior Scientific Officer, WARPO
A. M. Zoraf	Scientific Officer, WARPO
Shuvro Bhowmick	Scientific Officer, WARPO
Hridoy Roy	Assistant Professor, DChE, BUET & Co-investigator
Bimol Nath Roy	Research Assistant, DChE, BUET
Md Mahmud Kamal Bhuiyan	Research Assistant, DChE, BUET

1.6 Report Outlines

The first chapter includes the background of the research, objectives and organization of the report. This chapter describes the necessity of this research on national level. The research team composition is also described in this chapter.

The second chapter includes an extensive literature review which describes the concept of water quality analysis and different types of statistical approaches for water quality analysis. Moreover, an in-depth discussion on different statistical approaches for water quality analysis has been thoroughly analyzed.

The third chapter describes the basic theory and concept of basic statistics and linear algebra required for principal component analysis (PCA), factory analysis (FA), cluster analysis (CA), and time series analysis.

The fourth chapter comprises the theories and literature reviews of water quality index (WQI). Overview of general model structure and most commonly used WQI models which are discussed briefly along with their evaluation process.

The fifth chapter includes the approaches and methodology of the study. The methodology comprises of site selection, data collection, ArcGIS modeling, sample testing and analysis, PCA and WQI development.

The sixth chapter focuses on data collection and analysis, detailing the preliminary data collection conducted prior to the acquisition of equipment. It includes an overview of the site selection process, specifying the chosen sites and the rationale behind their selection. The chapter also covers primary data collection, highlighting the trends observed in each parameter. Additionally, statistical analyses such as box plots and MANOVA are employed to further interpret the data, providing insights into the variations and relationships between the collected parameters.

The seventh chapter focuses on the development of the Water Quality Index (WQI) model, outlining the key steps involved in its creation, including parameter selection, weight calculation, sub-indexing, and the formulation of the aggregating function. Once the WQI model is developed, the chapter proceeds with the evaluation of the WQI using the collected data, assessing the water quality across different sites and seasons.

The eighth chapter presents a comparative analysis of WQI values developed using datasets from 2019 and 2024, highlighting temporal changes in water quality across selected rivers. Additionally, this chapter includes the evaluation of the CCME WQI for 2024 to provide an alternative and robust assessment of surface water quality without relying on sub-indexing or parameter weighting.

The nineth chapter provides discussion and strategic recommendations to enhance the project's future success, including continuous monitoring, utilizing the developed Water Quality Index (WQI), and considering the implications of evolving legislative laws and water quality regulations. The tenth chapter presents the concluding remarks of the project. The eleventh chapter includes references.

The twelfth chapter of the report includes the appendices that provide supplementary information supporting the main content. Appendices A to G document various aspects of project coordination and activities undertaken throughout the study. Appendix H outlines the detailed timeline of project activities. Appendix I presents the original research proposal, while Appendix J contains the Memorandum of Agreement (MoA) outlining the contractual framework. Appendix K provides the geographical locations of the selected sampling sites. Appendices L and M include photographic documentation of field visits and sample collection activities, respectively.

Chapter 2: Water Quality Analysis (WQA): Theoretical and Statistical Approaches

2.1 What is Water Quality Analysis

Water quality analysis (WQA), also called hydrochemical analysis, is the assessment and evaluation of the physical, chemical, and biological properties of water to determine its overall quality and appropriateness for specific uses. It includes the collection, measurement, and interpretation of numerous parameters and indicators to evaluate the safety level of the water sources. Water quality analysis is crucial for a variety of reasons, including human health, environmental impact, regulatory compliance, industrial and agricultural processes, early prediction of potential pollutions, etc. Contaminants such as bacteria, viruses, heavy metals, pesticides, and chemical pollutants can have adverse effects on human health as well as industrial operations (Chowdhary et al., 2020). Water quality analysis helps to identify and quantify these contaminants to assess the potential risks associated with water consumption. Industrial wastewater discharges, agricultural runoff after pesticide and fertilizer usage, and improper municipal solid waste disposal can introduce toxic elements into water bodies, leading to ecological imbalances and danger to aquatic life (Garg et al., 2022). Industries and agricultural activities often require a significant amount of water usage. In industries, substandard water quality may cause product and safety hazards resulting in extreme economic loss. Water quality analysis is therefore crucial in these sectors to ensure that the water used is suitable for specific processes, such as manufacturing, irrigation, or livestock consumption. WQA helps identify any potential issues that could affect the efficiency of industrial processes or harm agricultural productivity. Governments and regulatory bodies enforce water quality regulations and guidelines to safeguard public health, the ecosystem, and the climate. Regular monitoring and analysis help identify potential violations and ensure maintaining water quality within acceptable limits. Improvement in water management and pollution control requires accurate predictions of water quality. A lack of information management contributes to the occurrence of numerous water pollution incidents. Therefore, the results of WQA help inform decision-making regarding water treatment, resource management, pollution control, and environmental conservation.

Water quality analysis starts with different laboratory testing and field measurements. Analytical techniques employed in water quality analysis include spectrophotometry (Hudson et al., 2008), chromatography (Kasiske et al., 1978), titration (Belle-Oudry, 2008), microbiological assays (Bonadonna et al., 2019), and molecular techniques (Girones et al., 2010). Figure 2.1 shows the general flow diagram of the WQA process (Roy, 2018).

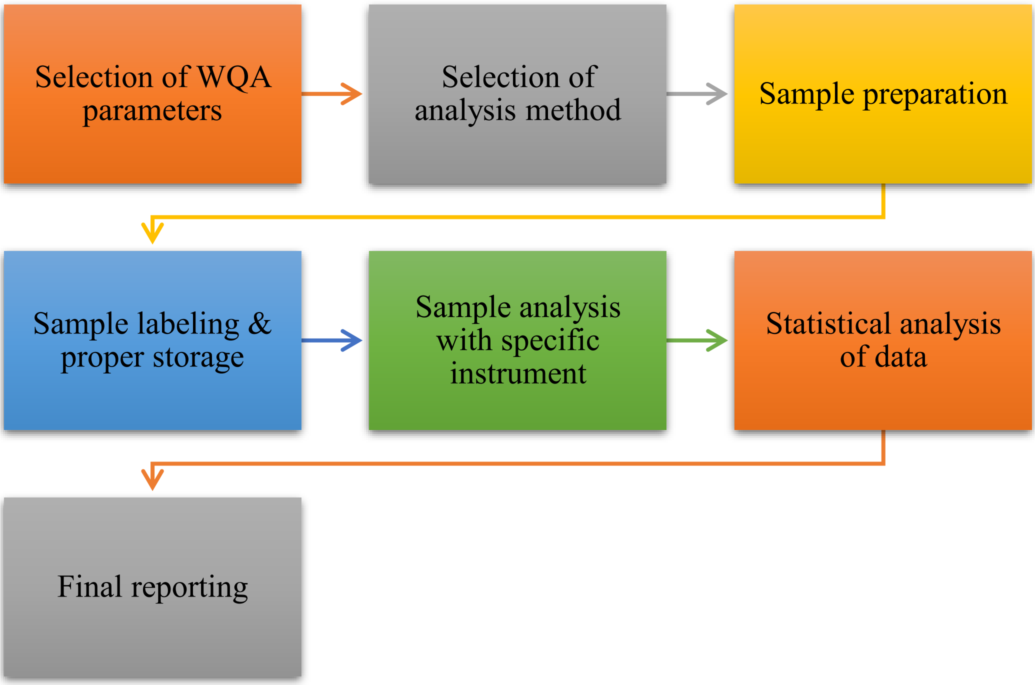


Figure 2.1: General flow diagram of water quality analysis process

Water quality indicators consist of physical (Duclos Alegue & Gnauck, 2006; Huey & Meyer, 2010; Pinto et al., 2013; Toming et al., 2016; Uthicke et al., 2012), chemical (Kannel et al., 2007; Maskooni et al., 2020; Sánchez et al., 2007; Ustaoğlu et al., 2020), and biological (Akay & Dalkıran, 2020; Al-Afify et al., 2019; Bigham et al., 2019; Summers & Summers, 2020) variables that provide explicit signals about the status and changes of the water system. The evaluation of these indicators is done through field monitoring of water bodies that provide important data for identifying water quality trends, providing relevant water authorities with information on water quality, and recommending future actions. Several essential freshwater quality indicators are provided in Table 2.1.

Table 2.1: Types of Water Quality Parameters

Physical Indicators	Chemical Indicators	Biological Indicators
Temperature	pH	Escherichia coli (E. coli)
Suspended Solids	Dissolved Oxygen	Coliforms
Turbidity	Electrical Conductivity	Biological Oxygen Demand (BOD)
Solar Radiation	Nutrients	Aquatic Macroinvertebrates
Color	Heavy Metal	

Physical indicators, e.g., temperature can be useful in measuring the degree of hotness or coldness of water, which can affect the solubility of substances and influence aquatic organisms (Bozorg-Haddad et al., 2021). Turbidity indicates the clarity or cloudiness of water caused by suspended particles, which can affect light penetration and restrain photosynthesis in the bottom layer of water bodies (Donald et al., 2019). Color can be useful as an indicator for the visual appearance of water, which can be influenced by natural substances or contaminants. The acidity or alkalinity of water is measured using a chemical indicator, such as pH. The solubility of minerals and the survival of aquatic organisms are both impacted by it, making it crucial. Conductivity, on the other hand, provides an estimate of its dissolved ion content and salinity. Dissolved Oxygen (DO) determines the amount of oxygen available in water to support aquatic life. Low DO concentrations may represent pollution or an excess of organic matter (Cooper, 1993). Certain levels of nitrogen, phosphorus, and other nutrients can promote excessive plant growth (eutrophication) and cause water quality issues. Biologically accessible nutrients and hazardous chemicals in excess can result in a number of disorders, e.g., toxic algal blooms, oxygen depletion, fish mortality, biodiversity loss, and the devastation of aquatic plant beds (Anderson et al., 2021). Again, effluents from municipal, agricultural, and industrial streams may contain harmful heavy metals such as lead (Pb), chromium (Cr), nickel (Ni), cadmium (Cd), iron (Fe), arsenic (As), and mercury (Hg) (Khan et al., 2023). Total Coliforms and Escherichia coli (E. coli) are biological indicators that suggest the potential presence of harmful pathogens in water. Aquatic Macroinvertebrates examine the diversity and abundance of small organisms living in water bodies that can provide insights into the overall ecological health and water quality. Globally, microbial indicators have been used to determine if a water body has been contaminated by human refuse. Typically, microorganisms found in high concentrations in human feces are utilized. In the United States, the most common indicators are total coliform, fecal coliform, Escherichia coli, and enterococci (Shibata et al., 2004).

It is particularly noteworthy that the specific indicators analyzed can differ based on the desired application of water (e.g., water for drinking, recreational purposes, industrial usage) and the regulatory standards applicable in a particular region. Therefore, water quality analysis often involves multiple indicators to obtain an accurate picture and extensive understanding of water quality and detect probable issues or origin of pollution.

2.2 Different Approaches for WQA

The most important natural resource is water. Accepting the significance and scarcity of resources required to meet biological needs and support all forms of economic and growth activity is of the utmost importance. One of the crucial concerns in the management of water resources is water quality. Water can be classified into two groups depending on the source: one is the groundwater, and another is the surface water. A variety of contaminants, including heavy metals, pesticides, fertilizers, toxic chemicals, and oils, may present as a result of domestic, industrial, and agricultural operations, which could result in the contamination of water (Omer & Omer, 2019). According to Swamee and Tyagee (2007) (Swamee & Tyagi, 2007), there are several metrics for each of the three major categories for assessing water quality - physical, chemical, and biological. By field monitoring of rivers, these three categories are evaluated. This information is used to identify patterns, notify water authorities about the quality of the water, and suggest future courses of action. The normal method of conducting this assessment is to consider the planned applications, human health, and natural water quality (Gazzaz et al., 2012; Pesce & Wunderlin, 2000). Water quality could be evaluated either spatially or temporally, as per Rosemond et al. (2009). However, measuring the various water quality indicators and comparing the measured values to threshold values are two essential components of water quality assessment. The threshold value denotes the maximum or minimum variable concentration that is safe for consumption by the general public.

Researchers have been focused on hydrochemical analysis in various ways for decades. The Piper diagram has been widely used to examine the groundwater facies in order to advance study on topics like disclosing the evolution of phreatic water and comprehending the hydrochemical properties as well as the mechanism of groundwater production (Wang et al., 2019; Yang et al., 2016). Traditionally, hydrochemical evaluation of groundwater was based on laboratory study, but for the last 20 years, geosciences research has been conducted in a variety of domains utilizing geospatial approaches, which have the advantage of monitoring and integrating many thematic levels with ease, precision, and on time (Ali & Ali, 2013; Ali & Pirasteh, 2004). Ahmad Ali et al. (2017) (Ali & Ali, 2018) analyzed the chemical variations in groundwater under diverse natural and anthropogenic activities, including the spatial distribution of hydrochemical parameters. The adequacy of groundwater quality for future domestic, agricultural, industrial, and drinking uses was also examined in this study. Basic techniques in hydrochemical research for compiling and presenting data on water quality include graphical and numerical interpretations. There are a large number of methods for categorization, correlation, analysis, and illustration (Sharma et al., 2012; A. K. Singh et al., 2008). In the past, water quality reports were generated using trend analysis utilizing a single variable on particular sample locations. Trend analysis can be done using both parametric and nonparametric tests.

Since the 1960s, water quality indices (WQIs) have been used as a technique to assess the condition of river water quality. WQI has grown in importance and popularity as a vital and widely used tool for evaluating the water quality of water bodies around the world, particularly rivers, due to its simplicity of use and scientific foundation. Since the introduction of the WQI idea, numerous scholars have created and developed numerous indices. Additionally, WQIs have been seen as a crucial component of broader environmental or natural resource indices such as the Environmental Performance Index (EPI 2010) (*2010 ENVIRONMENTAL PERFORMANCE INDEX*, n.d.). Canadian Council of Ministers of the Environment Water Quality Index (2001) and Oregon Index (Dhany Sutadian et al., n.d.) are the widely used indices for the purpose. The study of water quality has given considerable attention to the use of various multivariate statistical techniques, such as cluster analysis (CA), principal component analysis (PCA), factor analysis (FA), and discriminant analysis (DA), which can assist in the interpretation of complex data matrices to better understand the water quality and ecological status of the studied systems and allow the identification of potential factors/sources that influence water quality (Yang, Hou, et al., 2015; Yang, Zhang, et al., 2015)

2.3 Different Statistical Approaches for WQA

Water quality data usually involves a large number of measurements. These data serve as a foundation for plant operation, modeling the process, treatment planning, and economic assessments. In national and global contexts, there is a growing demand for methods to integrate several water quality-related variables into a single index. Globally, many scientific methods have been used to assess and monitor water quality. For detecting the origins of physical and chemical features and their associations, as well as disseminating information about the overall quality of the water, multivariate statistical techniques have recently gained popularity (Kazi et al., 2009). Multiple authors have incorporated water quality variables into indices, formally known as WQI. Increasingly, computational methods based on artificial intelligence have been applied to environmental problems with inherent uncertainties and subjectivities (Li et al., 2019). However, conducting different statistical analyses on the same data may significantly increase the statistical power and acceptance of the method.

Even though many WQIs have been designed, there is no widely accepted way to carry out the WQI development process globally (Dhany Sutadian et al., n.d.). Many researchers found it ambiguous to define WQI and convey it in a concise and unified manner. Due to the multifaceted nature of the elements or indicators that affect water quality and their wide range of value, this challenge became prominent. Since having several measurements in the data may affect the prediction and accuracy, popular multivariate statistical techniques, e.g., principal component analysis (PCA), factor analysis (FA), cluster analysis (CA), and time series analysis are widely used to group the important variables and are reviewed in this study.

2.3.1 Principal Component Analysis (PCA)

PCA is predominantly used to identify trends and patterns in a complex dataset by converting a set of potentially correlated observations into an entirely novel set of linearly uncorrelated variables, known as principal components. PCA is primarily employed for exploratory data analysis and the development of models for forecasting (Jollife & Cadima, 2016). It aims to explain as much variance as possible with the lowest possible number of principal components while maintaining a minimum information loss (Zou et al., 2006). When there are countless predictors in relation to the number of data, PCA has the potential to scale down the number of factors significantly (Chan et al., 2022).

In terms of mathematics, PCA consists of the following five main steps: (1) ensure that all of the measures are given the same weight in the statistical evaluation, defining the elements x_1, x_2, x_3,..x_p with zero means and unit variance, (2) estimating the covariance matrix C, (3) determining the eigenvalues λ_1, λ_2, λ_3,… λ_P and the associated eigenvectors a_1, a_2, a_3,..a_P, (4) discarding elements that merely contribute a minor amount to the variance in the datasets, and (5) developing the factor loading matrix to determine the principal parameters (Chauhan & Sharma, 2003).

PCA is the most basic form of multivariate analysis based on true eigenvectors (McGarigal et al., 2000). It has a close relation to factor analysis, which is related to canonical correlation analysis (CCA) (Matthew et al., 1994). PCA defines a new orthogonal coordinate system that optimally describes the variance in a single dataset, whereas CCA defines coordinate systems that optimally describe the cross-covariance between two datasets. Recent years have seen the application of the PCA model to a wide range of sustainability problems, such as the assessment of groundwater quality monitoring, pollution source determination in ambient soil and air, heavy metal detection, climate change analysis, etc. (Olsen et al., 2012; Ouyang, 2005).

Ouyang et al. (2006) analyzed river water quality using PCA to assess seasonal correlations along the main stem in Florida, United States (Ouyang et al., 2006). PCA and FA were conducted using statistical analysis system (SAS) software. Results from the analysis indicated that the importance of any water quality indicator, e.g., temperature, total alkalinity, color, and BOD depends significantly on the particular season, except dissolved organic content (DOC) and conductance. Bengraïne and Marhaba (2006) demonstrated the usage of PCA to identify the components linked to the variations in hydrochemistry using 19 independent parameters in the Passaic river, New Jersey (Bengraïne & Marhaba, 2003). PCA was used since the units or the variables differ substantially in this study. Moreover, to solve the issue of variables loading substantially on one or more of the axes, a varimax rotation was also carried out in that study.

2.3.2 Factor Analysis (FA)

FA is a statistical technique used to investigate the latent factors or dimensions underlying a data set. It is similar to PCA, but it concentrates on identifying the common factors that explain the correlations between observed variables, as opposed to maximizing variance. The FA can be expressed as the following equation (S. K. Singh et al., 2016).

$$Z_{ji} = a_{f1}f_{1i} + a_{f2}f_{2i} + a_{f3}f_{3i} + \cdots + a_{fm}f_{mi} + e_{fi}$$

Where Z_{ji} is the component score, a is the component loading, f is the factor score, e is the residual term accounting for errors or other sources of variation, i is the sample number and m is the total number of variables.

The primary objective of factor analysis is to reduce the contribution of insignificant variables in order to further consolidate the data structure generated by PCA (Wunderlin et al., 2001). FA assumes that the observed variables are influenced by a smaller number of latent factors, which are not explicitly observable but can be inferred from the intercorrelation patterns between the observed variables. The objective of factor analysis is to identify these latent factors and comprehend their contribution to the observed data. The applications of PCA and PFA in environmental management and protection research were greatly illuminated by recent works (Ouyang et al., 2006).

2.3.3 Cluster Analysis (CA)

Similar to factor analysis, cluster analysis also groups the data or variables into homogeneous clusters and after getting clusters, the correlation between homogeneous clusters can be identified. The hydrogeochemical parameters are first categorized using cluster analysis based on their average likeness. Typically, the analysis uses the squared Euclidean distance and Ward's linkage approach (Bu et al., 2020).

2.3.4 Time Series Analysis

Analyzing water quality data collected over different time domains, time series analysis is applied. A time series is a collection of data points that are usually recorded at regular intervals over a period of measurements. It examines patterns, trends, and seasonality in the data and can identify long-term or periodic changes. It is possible to analyze and interpret temporal patterns in water quality parameters using techniques such as autocorrelation analysis, moving averages, and seasonal decomposition.

The monthly variance of water quality standards was utilized by Parmar and Bhardwaj (2014) to compare statistical parameters, e.g., mean, median, mode, standard deviation, kurtosis, skewness, and coefficient of variation (Parmar & Bhardwaj, 2014). Tan et al. (2012) studied the least squats support vector machine LS-SVM algorithm to build a nonlinear time series forecasting model to predict water quality (Tan et al., 2012). Figure 2.2 shows the framework for the LS-SVM approach, which was adopted to have a high prediction accuracy and to be better suitable for small sample sizes in real-time water quality data forecasts.

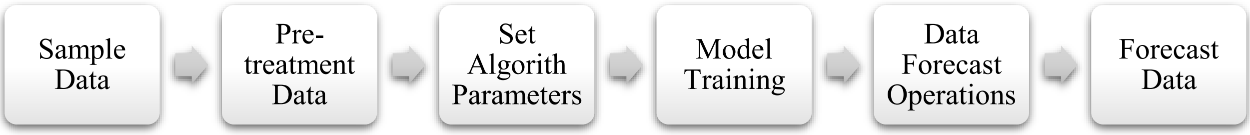


Figure 2.2: Implementation framework for prediction of water quality

2.4 Why is principal component analysis superior?

The quality of water has been analyzed by different researchers using various parameters. Quality, however, is a general phrase that is difficult to define using precise data. For instance, good quality water cannot simply be defined as having a pH value of 7.0 or higher (Mahapatra et al., 2012). Monitoring any potential increasing or declining trends in water quality is therefore essential to understand how the quality of the water is changing. In order to accomplish this, the responsible parties should be involved in implementing suitable monitoring programs that include all necessary criteria to identify the natural and anthropogenic mechanisms regulating water quality (Benkov et al., 2023). A better understanding of the condition of the environment can be obtained by characterizing the spatial variation and source allocation of water quality measures, which can also assist policymakers in setting priorities for sustainable water management (Huang et al., 2010). To accurately assess WQI, it is important to determine which parameters are appropriate. According to agency guidelines like the World Health Organization (WHO) (2006), physical, chemical, and biological indicators are typically used to determine the quality of water. These indicators include pH, electrical conductivity (EC), total dissolved solids (TDS), total suspended solids (TSS), hardness, turbidity, and contaminant concentrations.

PCA is a well-established statistical method for the investigation of water quality (Acquavita et al., 2015; Navarro et al., 2010; Platikanov et al., 2014) by analysis of obtained observed data. It is a statistical tool for data reduction that can be used to evaluate controls on groundwater composition, interpret observed relationships between variables, and produce simpler relationships that put insight into the underlying structure of the variables. It can also be used to aggregate the effects of many variables into a small subset of factors (Liu et al., 2003). By analyzing the primary sources of data variation, the latent factors (principal components) generated by the original variables (the water quality indicators) are used to interpret the data ("Principal Component Analysis for Special Types of Data," 2002). With the least amount of original data loss possible, PCA offers information on the important parameters (K. P. Singh et al., 2004). The original variables are combined in weighted linear combinations to form the principal components (PCs). PC gives data on the most important factors that describe the entire dataset and allows for data reduction with the least amount of original information lost (Filik Iscen et al., 2008).

In a recent study by Ersan Batuar et al. (2021) (Batur & Maktav, 2019) the findings of multiple linear regression (MLR), artificial neural network (ANN), and support vector machines (SVM) data mining methods were compared with those obtained using PCA-based response surface regression (RSR) approach for calculating surface water quality parameters. The investigation proved that the PCA-based RSR method is more accurate at estimating lake water quality parameters than MLR, ANN, and SVM data mining models. In 2008, T.G. Kazi et al. showed the origins or variables that cause changes in water quality were found using PCA (Kazi et al., 2009). In their studies of the regional variability of surface water quality and source apportionment, Shrestha and Kazama (Shrestha & Kazama, 2007), Huang et al. (Huang et al., 2010), and Juahir et al. (2011) (Juahir et al., 2011) divided the analyzed water bodies into three categories: High pollution site (HP), Sites with moderate pollution (MP) and low pollution (LP). PCA aids in the analysis of complicated data matrices so that the water quality and ecological status of the system under study can be better understood. These technologies make it easier to identify potential water quality influencing elements, and they can help with dependable water resource management as well as rapid pollution problem solutions (Adams et al., 2001).

Table 2.2: Comparison of different statistical approaches for WQA

Criteria	PCA	Factor Analysis	Cluster Analysis	Time Series Analysis
Objective	Maximize variance	Identify underlying factors	Identify groups/ cluster	Analyze data over time
Method	Linear transformation	Linear transformation	Grouping based on similarity	Analysis of temporal data
Use case	Dimensionality reduction	Identifying latent variables	Market segmentation	Forecasting, trend analysis
Strengths	Simplicity, effectiveness	Understand underlying structure	Discover natural groupings	Analyze temporal patterns
Limitations	Only linear relationships	Interpretation can be complex	Can be computationally intensive	Requires time dependent data

Chapter 3: Principal Component Analysis (PCA): Theory and Literature

3.1 Theory Behind PCA

For understanding PCA, basic knowledge of linear algebra and statistics is important (Ringnér, 2008). To understand the basic method of performing PCA, the following statistical concepts and some basic theories of linear algebra are discussed.

3.1.1 Basic Statistics Behind PCA

Understanding or having insights about three statistical variables that describe the trend in the dataset would help to get the idea of PCA. These three variables are – a) the mean, b) the variance and c) the covariance.

a) The mean: The mean defines the middle of the data or the average value per sample of the data. Mathematically,

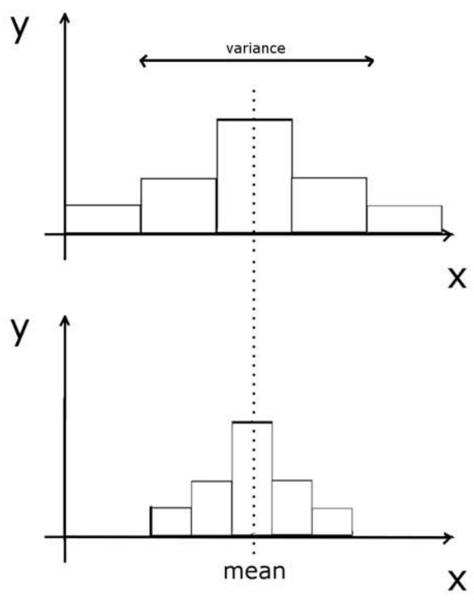
$$\bar{S} = \frac{1}{n}\sum_{i=1}^{n} S_i$$

Where n is the number of elements in the set S and $\bar{S}$ is the mean value of S.

b) The variance: Variance represents the spread of the data. Higher variance means a higher range of datasets. For the dataset S with n elements, variance can be expressed as –

$$Var(S) = \frac{1}{n-1}\sum_{i=1}^{n}(S_i - \bar{S})^2$$

Note that, $n-1$ is used for sample data while n is used for the whole population. The following illustration describes the variance in a graphical way. Both the dataset have the same mean but the dataset with higher variance is more spread.



c) The covariance: covariance defines the degree of codependence of two variables. If two variables are X and Y, then the covariance of X and Y is given by –

$$Cov(X,Y) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})$$

Now, for the covariance of X with itself,

$$Cov(X,X) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})(X_i - \bar{X}) = Var(X)$$

With these statistical insights, we proceed to the next section.

3.1.2 Basic Linear Algebra Behind PCA

Now that we know what the required statistical ideas are, we need to use them in our system. We have used the one-dimensional array or data to explain the statistical concepts. But in real life, datasets are made of different samples and parameters. Most of the time, 2D arrays are used which is also applicable for water quality assessments. There will be different samples that will be analyzed to determine different water quality parameters. Thus, we have to deal with 2D arrays or what we can call matrices. Suppose we have m number of parameters for each sample and n number of samples. If we build a dataset of all the values, then the dataset will look like the following –

$$D = \begin{bmatrix} D_{11} & D_{12} & \dots & \dots & D_{1n} \\ D_{21} & D_{22} & \dots & \dots & D_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ D_{m1} & D_{m2} & \dots & \dots & D_{mn} \end{bmatrix} = [D_{ij}]$$

where, D_{ij} represents the ith parameter of the jth sample and D is the dataset. Now if we think about applying the statistical concepts in this dataset for each parameter we will have,

$$\bar{D_\iota} = \frac{1}{n}\sum_{j=1}^{n} D_{ij}$$

where, $\bar{D_\iota}$ is the mean of ith parameter. The covariance of taking any two of the parameters can be expressed as –

$$Cov(D) = \frac{1}{n}DD^T$$

where, D^T is the transpose of matrix D. $Cov(D)$ is called the covariance matrix of D having a dimension of $m \times m$ and is given by,

$$Cov(D) = \begin{bmatrix} Cov(D_1, D_1) & Cov(D_1, D_2) & \dots & \dots & Cov(D_1, D_m) \\ Cov(D_2, D_1) & Cov(D_2, D_2) & \dots & \dots & Cov(D_2, D_m) \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ Cov(D_m, D_1) & Cov(D_m, D_1) & \dots & \dots & Cov(D_m, D_m) \end{bmatrix} = \left[Cov\left(D_p, D_q\right)\right]$$

where, $Cov\left(D_p, D_q\right)$ is the covariance of pth and qth parameter.

Now that we have implemented the statistical concepts in the matrix system, we need basic insights into eigenvectors and eigenvalues, the two most important concepts of linear algebra. Eigenvectors are those which does not change direction in any transformation. For any square matrix A, there will be –

$$Ax = \lambda x$$

Where λ is a scaler and x is a vector. Here, multiplying a matrix by a vector produces the same result as multiplying a scalar by that vector. The scaler value λ is called the eigenvalue and the vector x associated with λ is called the eigenvector. Different eigenvectors are found for different eigenvalues. Now, it can be written that,

$$Ax - \lambda x = 0$$

$$Ax - \lambda Ix = 0$$

where I is the identity matrix. Thus,

$$|A - \lambda I| = 0$$

and solving this equation for λ will determine the eigenvalues and using the first equation, eigenvectors can be determined. With these insights on statistics and linear algebra, we can proceed to perform PCA for any dataset.

3.1.3 Execution of PCA

As mentioned before, a matrix D with $m \times n$ dimension is taken. Performing PCA includes the following steps –

1. Determine the mean of each parameter/variable of the matrix D.

$$\overline{D_\iota} = \frac{1}{n}\sum_{j=1}^{n} D_{ij}$$

2. Subtract the corresponding mean of any parameter from each row. A new matrix D' will be found which will be called a shifted data matrix. This is done to standardize the data. D' can be expressed as,

$$D' = \begin{bmatrix} D_1 - \overline{D_1} \\ D_2 - \overline{D_2} \\ \dots \\ \dots \\ D_m - \overline{D_m} \end{bmatrix}$$

3. Calculate the covariance matrix, C for D'

$$C = \frac{1}{n} D' D'^T$$

4. Calculate the eigenvalues and eigenvectors of the covariance matrix, C
5. Check for length of eigenvector, eigenvector should be of unit length.

$$length = \sqrt{sum\ of\ the\ square\ of\ the\ elements\ of\ eigenvectros}$$

if $length \neq 1$, then rescale them with the following expression,

$$Scaled\ element = \frac{original\ element}{length}$$

6. Place the eigenvectors side by side with the eigenvector of the largest eigenvalue to the left and the descending.

$$W = [e_1 \quad e_2 \quad \dots \quad \dots]$$

where, e_1 is the largest eigenvalue.

7. Transpose the obtained matrix W and express the data in new rotated coordinate by,

$$D_{PCA} = W^T D'$$

The coordinate transformation that we generated with this method has rotated the axes so that the first coordinate axis lines up with the data in such a way that most of the variation is in that coordinate. The data were correlated in the original coordinates, but they are not correlated in the new rotated coordinates. The aim of PCA is to derive a new set of coordinates for the data that are uncorrelated and that are in the order of the degree of variation in that coordinate. How many of the new coordinates (also called components) are kept is an arbitrary choice; however, the intention is to keep only enough components to capture the essence of the data. A typical method of selecting the components to keep is to sum all of the eigenvalues and then keep only those components with the largest eigenvalues, which sum up to no less than 90% of the total. This is done because the larger the eigenvalue, the greater the amount of variation of the data in the direction of the corresponding eigenvector. In PCA, we choose to keep only those components that carry most of the variation of the data.

Chapter 4: Water Quality Index: Theory and Literature

4.1 Overview of Water Quality Index Modeling

Modeling of WQI can be categorized into four main steps (Abrahão et al., 2007; Lumb, Sharma, Bibeault, et al., 2011; Sutadian et al., 2018a), namely –

1. Selection of water quality parameters that are most important for evaluating water quality
2. Sub-indexing the water quality parameters for making a unitless comparison between the parameters
3. Weighting the water quality parameters based on the significance of the water quality
4. Using an aggregating function to determine the WQI by aggregating the sub-indices into a single value

Although some models have been developed with fewer steps, most of the models considered these four steps. The general model structure for developing WQI models is illustrated in Figure 4.1.

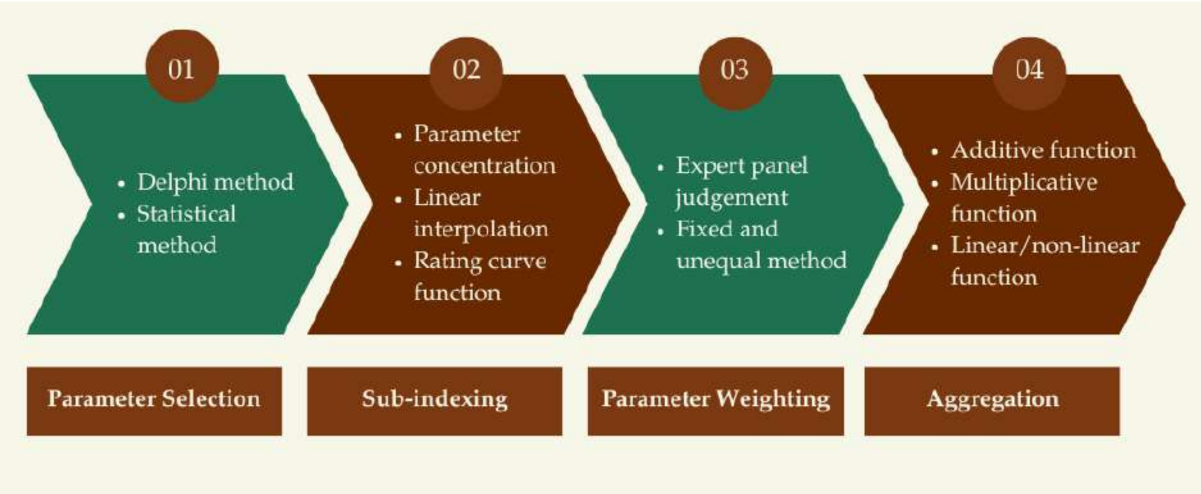


Figure 4.1: General model structure of the WQI model

4.1.1 Parameter Selection

Parameter selection is a rudimentary step for modeling WQI for any given instance. Parameters selection can be executed based on the availability of data, expert opinion or the environmental significance of a water quality parameter. WQI has been modeled with as high as 26 parameters such as in the Bascaron Index and Dojildo Index (Koçer & Sevgili, 2014; M. G. Uddin et al., 2021), but most of the models deal with 8 to 11 parameters. However, models such as CCME-WQI were developed considering only 4 parameters (CCME WATER QUALITY INDEX USER’S MANUAL 2017 UPDATE, 2017). Many of the researchers did not include some parameters, for example, suspended solids, microbiological contamination in their model because of high analytical cost and limited analytical laboratory facilities (Ma et al., 2020;

Naubi et al., 2016). Galal et al. reported that in most of the developed WQI models the Delphi technique was followed to select the water quality parameters based on gathering expert opinions by interviewing or surveying (M. G. Uddin et al., 2021). The Delphi technique, developed by Dalkey and Helmer in the 1950s is a widely acknowledged method for selecting important opinions of one's interest from a vast set of options available generally by using questionnaires to collect data from an expert panel (Hsu & Sandford, 2019). However, some critics suggest that this technique produces data uncertainty and low model accuracy (Goodman, 1987; Linstone, 1985). Generally, there are no specific rules for selecting water quality parameters to include in the WQI modeling. Recently, statistical approaches have been taken to select water quality parameters and in the field of water quality assessments (Benkov et al., 2023; Tharmar et al., 2022). A statistical approach for selecting water quality parameters that will be followed in this study is principal component analysis (PCA). Recent studies on water quality assessment suggest that PCA can be a better tool to use for evaluating surface water quality (G. Uddin et al., 2022). As there are a copious amount of variables available for evaluating water quality, the relationship among the variables is supposed to give better interpretations than the variables themselves. Also, PCA can be used to find out the dominant variables and their impact on the water quality (Tharmar et al., 2022). By this method, a minimum number of useful parameters can be selected to include in the WQI model with higher efficiency compared to the number of parameters. Lowering the number of parameters is important due to the time consumption for determining all the parameters analytically and economically.

4.1.2 Sub-indexing

Parameters for evaluating water quality assessment are immensely disparate in the unit. Thus, sub-indexing is done to get rid of the units and make all the parameters comparable in one format. Most of the model uses standard guideline values to achieve sub-indices (Abbasi & Abbasi, 2012c; S. M. Liou et al., 2004). However, some of the models, for example, CCME-WQI skipped this step and worked with the aggregating function directly (CCME WATER QUALITY INDEX USER'S MANUAL 2017 UPDATE, 2017).

There are several ways for sub-indexing. The most used sub-indexing method is the linear interpolated function where parameters are converted according to their water quality standard value from 0 to 100 (Effendi et al., 2015a; Lobato et al., 2015a) with the help of any given equation or graphical representation. However, models such as the Horton index and Said index use the measured parameter concentration directly as the sub-index which is the simplest way to do so. Another way of sub-indexing is using rated curving functions used in MRWQI and OWQI (Gazzaz et al., 2012; Hasan et al., 2015). This method is also based on the water quality parameter standard guideline values imposed by the legal authorities. Rather than taking a linear relation, this method takes logarithmic or non-linear regression to achieve sub-indices.

4.1.3 Parameter Weighting

The main purpose of this step is to apply the relative importance of the water quality parameters (Sarkar & Abbasi, 2006b). Most of the models uses an unequal weighting technique where in some of the models, the total sum of all the parameter weight is equal to 1. Some models, for example – the Horton index and Bascaran Index use integer value of the parameter weight causing the total sum to be greater than 1. A model such as OWQI uses equal weight for all the parameters. However, the CCME-WQI model does not require parameter weighting to evaluate WQI. The analytical hierarchy process (AHP) method, developed by Thomas Satty, is a decision-making technique that can be an effective method for evaluating parameter weighting (M. G. Uddin et al., 2021). This method can reduce uncertainty and give a better accuracy of the weighting process which can be found with AHP (Sutadian et al., 2017).

4.1.4 Aggregating Function

The final process of modeling WQI is the aggregating process. It is used to aggregate all the sub-indices into a single index for evaluating water quality (Dhany Sutadian et al., n.d.). This single index is called the WQI. There are different aggregating functions to achieve this step. These are discussed below –

i. Additive function: The additive aggregation function is one of the simplest ways for aggregating the sub-indices used in several models. The additive function can be expressed as –

$$WQI = \sum_{i=1}^{n} w_i \, S_i$$

Where, S_i is the sub-index of the ith parameter and w_i is the parameter weight value. This method has a major issue called 'eclipsing,' which means that the final index value does not represent the whole scenario (Dhany Sutadian et al., n.d.). Rather, it is dominated by the larger values of the sub-indices and neglect relatively lower values. There are variations in the additive function. Some models used the modified additive function which can be expressed as (Bordalo et al., 2006; Carvalho et al., 2011) –

$$WQI = \frac{1}{100}(\sum_{i=1}^{n} S_i)^2$$

$$WQI = \frac{1}{100}(\sum_{i=1}^{n} S_i \, w_i)^2$$

Where, S_i is the sub-index of the ith parameter and w_i is the parameter weight value. Another modified form of additive function is proposed by Bascaron (Dhany Sutadian et al., n.d.) and adopted and modified by some other models (Pesce & Wunderlin, 2000; Sánchez et al., 2007). In this function, the total value of final aggregation is divided by the total weight, i.e., the sum of the weights of the parameters. This modified function is expressed as –

$$WQI = \frac{\sum_{i=1}^{n} C_i P_i}{\sum_{i=1}^{n} P_i}$$

Where, C_i is the sub-index of the i^{th} parameter called the normalization factor and P_i is the parameter weight value. If a total weight of 1 is taken, then this equation becomes identical to the primary additive function.

ii. Multiplicative function: Another common aggregation function suggested by Brown and later included in the different models (Almeida et al., 2012; Devendra Swaroop Bhargave & Asce, 1985) is the multiplicative function which can be expressed as follows –

$$WQI = \sum_{i=1}^{n} S_i^{w_i}$$

Ambiguity arises in this method when the parameter weight is close to zero (S. M. Liou et al., 2004). Then this method fails to provide the actual water quality.

iii. Minimum operator function: The minimum aggregating method considers the minimum sub-index as the final WQI.

$$WQI = Min(S_1, S_2, S_3, \dots \dots \dots, S_n)$$

This method omits the eclipsing and ambiguity problem but fails to provide an overall composite picture of water quality (Swamee & Tyagi, 2000).

iv. Combined function: Liou et al. proposed a combined model for aggregating functions to avoid eclipsing and ambiguity problems (S. M. Liou et al., 2004). This aggregating function was proposed for evaluating river status and expressed as –

$$WQI = C_{temp} C_{pH} C_{tox} \left[\left(\sum_{i=1}^{l} I_i w_i \right) \times \left(\sum_{j=1}^{m} I_i w_i \right) \times \left(\sum_{k=1}^{n} I_i w_i \right) \right]^{1/3}$$

Here, C_{temp}, C_{pH} and C_{tox} are the three scaling factors which are the sub-indices for temperature, pH and toxicity, respectively. In this study, the results from the principal component analysis were used and the results were clustered into three groups. At first, the variables found from PCA are aggregated in the additive method for each principal component. Then the geometric mean is taken and finally, WQI is determined by a multiplicative function with three scaling factors and aggregated indices.

v. The square root of the harmonic mean function: Another proposed aggregating function is the square root of the harmonic mean function (Cude, 2001; Dojlido et al., 1994). This function can be expressed as –

$$WQI = \sqrt{\frac{n}{\sum_{i=1}^{n} \frac{1}{S_i^2}}}$$

Where, S_i is the sub-index of i[th] parameter. A study by Swamee and Tyagi (Swamee & Tyagi, 2000) stated that this aggregation method might cause a problem where all the sub-indices are acceptable but the overall index is not.

vi. Unique linear/non-linear aggregating function: Said et al. suggested a new aggregating function that deals directly with the parameter value of some preselected parameters and sub-indexing is not required (Said et al., 2004). The WQI equation is proposed as –

$$WQI = \frac{(DO)^{1.5}}{(3.8)^{TP}(Turb)^{0.15}(15)^{FCol/10000} + 0.14(SC)^{0.5}}$$

Where DO is the dissolved oxygen (%), T is turbidity (NTU), TP is total phosphate (mg/L), FCol is fecal coliform bacteria (counts/100 mL) and SC is the specific conductivity (MS/cm). The major limitation of this method is the study is based on a specific region and might cause inaccuracy for other regions.

4.2 Existing WQI Models

Galal et al. studied 107 cases and stated that almost 85% of the cases used seven models naming The Horton index, National Sanitation Foundation WQI (NSF-WQI), Scottish Research Development Department Index (SRDD index), Canadian Council of Ministers of the Environment WQI (CCME-WQI), Bascaron index (BWQI), Fuzzy Interface System (FIS) based index, Malaysian WQI (MWQI) (M. G. Uddin et al., 2021). Each of these models has a minimum of four applications, with the highest 36 applications of CCME-WQI and NSF-WQI at the second ranking based on the application on the case studies with a number of 18 applications. Another mentionable WQI model is the West Java WQI (WJ-WQI) model, not because of the number of applications but because it is one of the most recent WQI models developed considering the reduction of uncertainty of the existing models (Sutadian et al., 2018a). These above-mentioned models are briefly described in the following sections.

4.2.1 The Horton Index

Horton established the very first WQI model in 1965 for defining the water quality of rivers (Gupta & Gupta, 2021; Horton, 1965; Mukate et al., 2019; Tirkey et al., 2013). This model includes the four general steps for WQI modeling. Parameter selection was based on the environmental impact, data reliability and relative influence on other parameters (Abbasi & Abbasi, 2012a). Water quality parameters such as DO, pH, coliforms, specific conductance, carbon chloroform extract, and alkalinity (based on percent entry of population upstream served by treatment) were considered (Shah & Joshi, 2017). In this model, sub-indexing was done using a linear scaling function and the values ranged from 0 to 100. Where 0 represents the worst quality and 100 represents the excellent quality (Horton, 1965). Parameter weighting was based on the Delphi method. The expert panel suggested weighting values from 1 to 4 for different parameters.

An additive function is used as the final aggregating equation for determining the WQI value, expressed as the following –

$$WQI = \left[\frac{\sum_{i=1}^{8} w_i s_i}{\sum_{i=1}^{8} w_i}\right] \times m_1 \times m_2$$

Where, m_1 and m_2 are the coefficient of temperature and obvious pollution, respectively. Values of m_1 and m_2 are either 0.5 or 1.0 based on the temperature and presence of emission.

4.2.2 National Sanitation Foundation (NSF) WQI

Brown modified the Horton model and developed the NSF-WQI model (Abrahão et al., 2007; Lumb, Sharma, Bibeault, et al., 2011). Parameter selection was done using the Delphi technique (Ewaid, 2017; Rocha et al., 2015) and considered 11 different parameters of five different categories namely physical, chemical, microbiological, nutrient and toxic parameters (Dhany Sutadian et al., n.d.; Lumb, Sharma, Bibeault, et al., 2011). Linear scaling was used for sub-indexing and the values ranged from 0 to 1. 1 is considered to be in the recommended guideline (M. G. Uddin et al., 2021). The Delphi technique is also followed to evaluate the parameter weighting with an expert panel. Unequal weighting values are used, and the total sum of the weighting is 1. However, later applications of this model have used some modified weight values for some parameters (Noori et al., 2019; Tomas et al., 2017). And for the final step, the original model was developed with an additive aggregating function but later in 1973, Brown et al. suggested the multiplicative function to aggregate the sub-indices into an overall index (Brown et al., 1973; M. G. Uddin et al., 2021).

The proposed method for comparing the water quality of various water sources is based upon nine water quality parameters such as temperature, pH, turbidity, fecal coliform, dissolved oxygen, biochemical oxygen demand, total phosphates, nitrates and total solids. The water quality data are recorded and transferred to a weighting curve chart, where a numerical value of Q_i is obtained. The mathematical expression for NSF WQI is given by

$$WQI = \sum_{i=1}^{n} Q_i W_i$$

Where, Q_i = sub-index for ith water quality parameter and W_i = weight associated with ith parameter and n is the number of parameters. The model was first proposed with 9 water quality parameters. These are – DO, pH, BOD, temperature, total phosphate, nitrate, turbidity, total solids and fecal coliform. With that, the model proposed a specific weighting value for each parameters. Later, researchers modified the number of parameters and changed the parameters. They also defined their own parameter weight values. Table 4.1 shows a brief summary of the parameter weight values used by different researchers.

Table 4.1: Different parameter weight value used for NSF WQI model in different studies

WQ Parameter	Model Recommended Weight, W_i	Weight value defined in different studies					
		(Effendi et al., 2015a)	(Babaei Semiromi et al., 2011)	(Shah & Joshi, 2017)	(Hoseinzadeh et al., 2015)	(Ewaid, 2017)	(Tomas et al., 2017)
DO	0.17	0.20	0.129	0.17	0.17	0.17	0.2857
pH	0.11	0.11	0.133	0.11	0.12	0.11	0.0714
BOD	0.11	0.13	-	0.11	0.10	0.11	0.2142
Temperature	0.10	0.12	-	0.10	0.10	0.10	-
Total Phosphate	0.10	0.12	-	0.10	0.10	0.10	0.1429
Nitrate	0.10	0.12	0.128	0.10	0.10	0.10	0.1429
Turbidity	0.08	0.10	0.155	0.08	0.08	0.08	-
Total Solids	0.07	0.08	0.100	0.07	0.08	0.07	-
Fecal Coliform	0.16	-	0.182	0.16	0.15	0.16	-
Total Odor Number	-	-	-	-	-	-	0.1429
Total Suspended Solids	-	-	0.173	-	-	-	-

4.2.3 The Scottish Research Development Department (SRDD) Index

The SRDD-WQI model was developed by the Scottish Research Development Department and was used to assess surface water quality from different tropical-sub-tropical countries such as – Iran, Romania, Portugal (Carvalho et al., 2011; Dadolahi-Sohrab et al., 2012; IONUŞ, 2010). In Eastern Thailand, a modified version SRDD model has been used for evaluating water quality (Bordalo et al., 2006). All three processes, parameter selection, sub-indexing, and parameter weighting, were conducted using the Delphi technique (M. G. Uddin et al., 2021). The SRDD model used the following modified additive aggregating function based on the NSF-WQI model (Lumb, Sharma, Bibeault, et al., 2011):

$$SRDDWQI = \frac{1}{100}\left(\sum_{i=1}^{n} S_i w_i\right)^2$$

4.2.4 Canadian Council of Ministers of Environment (CCME) WQI

The CCME-WQI model is a modified version of the British Columbia Water Quality Index (BCWQI) model (Abbasi & Abbasi, 2012d). The CCME-WQI has been in application with a wide range of surface water bodies (Md. G. Uddin et al., 2017). The CCME-WQI requires a minimum of four parameters to determine the water quality index; parameters are not specified and decided by the users (CCME WATER QUALITY INDEX USER'S MANUAL 2017 UPDATE, 2017). This model does not include parameter sub-indexing and weighting. After user-defined parameters are selected the model directly uses the aggregating function. In this model, 3 factors F_1, F_2 and F_3 namely, 'scope,' 'frequency' and 'amplitude' respectively are calculated before aggregating the parameters. These factors are defined and expressed as follows –

F_1 = percent parameters that do not meet the specified objective

$$F_1 = \frac{number\ of\ failed\ parameters}{number\ of\ total\ parameters} \times 100$$

F_2 = percent individual test values that do not meet the specified objective value

$$F_2 = \frac{number\ of\ failed\ tests}{number\ of\ total\ tests} \times 100$$

F_3 = amount by which test values test values failed to meet their objective and this factor is calculated in three steps –

a. The number of times by which any individual parameter deviates from the objective is called excursion and is calculated using the following expressions –

$$excursion_i = \left(\frac{Failed\ Test\ Value_i}{Objective_i}\right) - 1\text{, when test values exceed the objective}$$

$$excursion_i = \left(\frac{Objective_i}{Failed\ Test\ Value_i}\right) - 1\text{, when test values fall below the objective}$$

b. The total extent of deviation is calculated by summing the individual excursions and dividing by the total number of tests

$$\text{the normalized sum of excursion, } nse = \frac{\sum_{i=1}^{n} excursion_i}{number\ \ of\ total\ tests}$$

c. Finally, an asymptotic function is used to calculate the factor F_3 to yield a range from 0 to 100 with the following expression –

$$F_3 = \frac{nse}{0.01\,nse + 0.01}$$

Finally, the function used for aggregation is expressed as:

$$CCMEWQI = 100 - \left[\frac{\sqrt{F_1^2 + F_2^2 + F_3^2}}{1.732}\right]$$

4.2.5 Bascaron Index (BWQI)

Bascaron developed this model based on the Spanish water quality guidelines in 1979 (Sun et al., 2016). Many South American countries and some southern Asian regions adopted the Bascaron Index and some other countries developed a modified indexing model based on the Bascarons WQI model (M. G. Uddin et al., 2021). BWQI model was developed with considering 26 parameters which is the highest number of parameters considered for developing any WQI model (Nong et al., 2020). Linear sub-indexing is applied based on the local water quality guidelines and a value ranging from 0 to 100 is developed against each of the parameters (Kannel et al., 2007; Pesce & Wunderlin, 2000). Unequal weight values ranging from 1 to 4 are used. Two modified additive functions were used to aggregate the sub-indices. One is to determine the objective index value and the other one is to determine the subjective index value. Expression for these two aggregating functions are as follows –

$$BWQI_{Obj} = \frac{\sum_{i=1}^{n} w_i S_i}{\sum w_i}$$

$$BWQI_{Sub} = k \times \frac{\sum_{i=1}^{n} w_i S_i}{\sum w_i}$$

The fundamental difference in these two equations is the constant, k which can be defined as the constant of visual assessment. values of k are achieved with a visual inspection of the river (Pesce & Wunderlin, 2000) and one of the following is taken as the value for k –

$k = 1.00$; clear water without apparent contamination of natural solids suspended.

$k = 0.75$; light contaminated water, indicated by light non-natural colour, foam, light turbidity for no natural reason.

$k = 0.50$; contaminated water, indicated by non-natural colour, light to moderate odour, high turbidity (non-natural), suspended organic solids, etc.

$k = 0.25$; highly contaminated water, indicated by blackish colour, hard odour, visible fermentation, etc.

4.2.6 Fuzzy Interface System (FIS) Based Index

Fuzzy logic has been a popular tool for decision-making since the 1960s and many researchers have used fuzzy interface systems for environmental assessments (Peche & Rodríguez, 2012). Several studies included FIS to develop WQI models in recent years (Lermontov et al., 2009; F. Lu et al., 2014; Xia & Chen, 2015). The FIS-based WQI model also uses four steps, and these are analogous to the general steps of developing WQI models (Lermontov et al., 2009). These steps are a) fuzzy sets and membership function – related to the parameter selection and based on the correlation studies; b) fuzzy set operations – normalizing the water quality parameters with FIS; c) fuzzy logic – weight values are determined with Fuzzy logic function; and d) interference rules – aggregating the water quality parameters using a range of fuzzy interference rules.

4.2.7 Malaysian Index (MWQI)

Department of Environment, Malaysia developed a WQI model in 1974 to evaluate the surface water quality based on the national water quality guidelines of Malaysia (Gazzaz et al., 2012). Six water quality parameters were selected for developing the model based on an expert panel. For sub-indexing, unique quality function curves were developed for each parameter. Parameter weighting was also done using the expert panel's opinion and unequal weighting was applied (Khuan et al., 2002). This model also uses the simple additive function for aggregating the sub-indices into the final WQI value. They classified the water quality based on the adopted WQI. Table 4.2 shows the water quality classification and Table 4.3 shows the WQI classification.

4.2.8 West Java Index (WJ-WQI)

This is the most recent WQI model developed in literature as per Sutadin et al. (2018) (Khuan et al., 2002). This model has all the four general steps of developing WQI models. This model was developed focusing on reducing the uncertainty of the previously existing WQI models. Parameter selection includes two sub-steps – a) using statistical assessment to evaluate parameter redundancy and b) identifying common parameters across the sampling stations. A linear scaling interpolation function is introduced to determine the sub-indices based on the maximum and minimum allowable guideline values of different water quality parameters. Parameter weighting was conducted with an expert panel and the panel's opinion was evaluated using the Analytical Hierarchy Process (AHP). Finally, the model uses the same multiplicative aggregating function used in the NSF-WQI model.

4.2.9 WQI Evaluation

The following table shows the evaluation classes recommended by different models.

Table 4.2: Different WQI Model Evaluation Classes

WQI Name	No. of Parameters Used and Evaluation Classes	WQI Range	Remarks on Water Quality
Horton Index	8 Parameters 5 Classes	91-100	Very Good
		71-90	Good
		51-70	Poor
		31-50	Bad
		0-30	Very Bad
NSF-WQI	11 Parameters 5 Classes	90-100	Excellent
		70-89	Good
		50-69	Medium
		25-49	Bad
		0-24	Very Bad
SRDD Index	10 Parameters 7 Classes	90-100	Clean
		80-89	Good
		70-79	Good Without Treatment
		40-69	Tolerable
		30-39	Polluted
		20-29	Several Polluted
		0-19	Piggery Water
CCME-WQI	4 Parameters 5 Classes	95-100	Excellent
		80-94	Good
		65-79	Fair
		45-64	Marginal
		0-44	Poor
BWQI	26 Parameters 5 Classes	90-100	Excellent
		70-90	Good
		50-70	Medium
		25-50	Bad
		0-25	Very Bad
MWQI	6 Parameters 3 Classes	81-100	Clean
		60-80	Slightly Polluted
		0-59	Polluted
WJ-WQI	13 Parameters 5 Classes	90-100	Excellent
		75-90	Good
		50-75	Fair
		25-50	Marginal
		5-25	Poor

Table 4.3: WQI parameters used in different WQI models

WQ Parameters	Existing WQI Models						
	Horton Index	NSF-WQI	SRDD Index	CCME-WQI	BWQI	MWQI	WJ-WQI
Temperature		√	√	Any Four Selected By User	√		√
pH	√	√	√		√	√	
Electrical Conductivity (EC)	√		√		√		
Total Solids (TS)		√					√
Total Dissolved Solids (TDS)	√						
Dissolved Oxygen (DO)	√	√	√		√	√	√
Biochemical Oxygen Demand (BOD)		√	√		√	√	
Total Suspended Solids (TSS)			√			√	
Chemical Oxygen Demand (COD)						√	√
Chloride (Cl^-)	√				√		√
Total Coliform (TC)					√		
Nitrate (NO_3^-)		√			√		√
Phosphate (PO_4^{3-})		√	√				√
Ammonia (NH_3)			√		√	√	
No. of Other Parameters	3	4	2		17	0	6

Chapter 5: Approach and Methodology

5.1 Introduction

For establishing the water quality index of Dhaka city-based rivers, samples will be collected bi-monthly from four rivers: Buriganga, Turag, Balu and Shitalakhya. Samples will be collected from 6 monitoring sites located on the selected rivers and 14 water quality parameters will be analyzed for each sample. After data collection, spatial interpolation of the sampling sites and parameter concentrations will be studied using ArcGIS software.

To achieve the main objective, the principal component analysis will be conducted using the sampling data and based on the result, i.e., based on the principal components, the WQI model will be developed by using a modified expression of the NSF-WQI model.

5.2 Site Selection

Site selection is mainly based on the Department of Environment's (DoE) sample collection points as well as on-site field visits. The Department of Environment has established a comprehensive monitoring network to diligently track surface water quality and this network entails collecting monthly samples from specific sampling stations strategically placed along rivers that are part of the network. Site selection will be done for all four rivers.

5.3 Data Collection

5.3.1 Primary Data

The methodology starts with a preliminary analysis of the surface water quality variables for the evaluation of temporal and spatial variations and the interpretation of the concentration of pollutants. This analysis helps to identify the most polluted segments of the four rivers and their adjacent streams, as well as the quality variables with the best and worst performances. This is a valuable step towards designing the more complex and rigorous procedure that we describe in the succeeding paragraphs. Initially, we would select 9 numbers of monitoring sites for each river based on the literature review of previous reports of WARPO, DoE and IWM, field surveys and analysis through the ArcGIS software. The sample collection would be done for each of the four seasons- pre-monsoon, monsoon season, post-monsoon and dry season. Table 5.1 shows the different parameters that will be considered for this study.

Table 5.1: Different considerable parameters and analyzing instruments/methods

Sl. No.	Parameter	Unit	Standard Protocol
1	Temperature	°C	APHA-2550
2	pH	-	ASTM-D-1293
3	Electrical Conductivity (EC)	mS/cm	Glass Electrode
4	Turbidity	NTU	-
5	Total Solids (TS)	mg/L	APHA-2540 B
6	Total Dissolved Solids (TDS)	mg/L	APHA-2540 C
7	Dissolved Oxygen (DO)	mg/L	Glass Electrode
8	Biochemical Oxygen Demand (BOD)	mg/L	APHA-5210 B
9	Total Suspended Solids (TSS)	mg/L	APHA-2540 D
10	Chemical Oxygen Demand (COD)	mg/L	APHA-5220 D
11	Chloride (Cl^-)	mg/L	ASTM-E-201
12	Total Coliform (TC)	MPN/100mL	US-EPA
13	E. Coli	MPN/100mL	US-EPA
14	Nitrate (NO_3^-)	mg/L	APHA-4500- NO_3^- B
15	Phosphate (PO_4^{3-})	mg/L	Spectrophotometric
16	Ammonia (NH_3)	mg/L	APHA-4500- NH_3 B

5.3.2 Secondary Data

In this methodology a secondary dataset of different water quality parameters for four Dhaka-based rivers will be collected from the Department of Environment (DoE) to establish a water quality index model with Principal Component Analysis (PCA). The water quality parameters available in that dataset may differ from the primary dataset but using this secondary data will help us to identify best practices and common pitfalls and to get practical insights of this WQI model development process.

5.4 ArcGIS Modeling

GIS is an indispensable tool for natural resource management, particularly in the areas of land use planning, animal habitat analysis, and natural hazard assessment to name a few of its many applications (Nandini et al., 2007). Although statistical surfaces, in the sense that they are defined by cartographers, do not exist in the same way that land does, it is possible to conceptualize them in the same way. For the preparation of statistical surfaces, the spatial interpolation process in GIS is commonly utilized (Wu et al., 2019). Given the impossibility of collecting field data at every location in the study area, spatial interpolation methods will be used to extrapolate information from sampled locations to locations where it could not be collected. Aiming toward the creation of maps depicting the spatial distribution of water quality parameters, ArcMap 10 GIS software would be used to correlate the water quality to the sampling locations (Nath et al., 2018). There are various spatial interpolation methods available. The inverse distance weighting (IDW) method will be used in this case (G. Y. Lu & Wong, 2008). By using the spatial interpolation method, maps depicting the spatial distribution of the water quality parameters with specified sites will be created. Identifying variations in the accumulations of various parameters in the surface water of the research area was performed.

5.5 Principal Component Analysis

A statistical approach, PCA, will be used to analyze the water quality of Dhaka-based four rivers (Buriganga, Turag, Balu and Shitalakshya). The developed method can be easily applied to other river basins suffering from similar environmental problems. The developed method can rigorously reflect the trends and patterns in water quality and is easy to apply and interpret. Among various techniques, PCA has received interest in water quality analysis. PCA can be used to identify the critical parameters that severely affect the water quality of a specific water body, which would significantly decrease the load of data for further data processing.

PCA will be utilized to reduce the huge amount of data into some meaningful, simpler data that will be further used to find out the prominent factors that are responsible for the variation of the water quality characteristics. To achieve this, R language will be used for data analysis. And software MS Excel and RStudio 2023.9.1.0 will be used.

R is a programming language and environment specifically designed for statistical computing and graphics. It is an open-source software that provides a wide variety of statistical and graphical techniques, making it a powerful tool for data analysis, manipulation, and visualization. R has a large and active community of users and developers, contributing to its extensive library of packages for diverse statistical methods. Key features of R includes – data analysis, statistical modeling, data visualization, data manipulation and programming. RStudio is an integrated development environment (IDE) for R that enhances the R programming experience. It provides a user-friendly interface with features tailored to R development, making it easier to write, debug, and execute R code. Community version of RStudio will be utilized in this research project. RStudio have several advantages on the user-friendly interface such as – script editor, console, data viewer, plots and charts, package management etc. which can be handled in one window altogether making it highly suitable for data analysis.

PCA for water quality parameters using R would be achieved with the following steps. Required steps for performing PCA using R is demonstrated in the Figure 5.1.

1. Setting up the working directory in the computer which is selecting the folder where the data files and analysis outcomes will be stored.
2. Data collection and data importing which includes saving data into a .csv (comma separated value) file and importing these data into a data frame for further analysis.
3. Data exploration and processing, examining the structure of the dataset and imputing missing values or removing outliers.
4. Principal component analysis which can be done using different built-in functions such as prcomp() or princomp() with specified function variables for data standardization including scaling and centering.
5. Interpretation of the PCA includes the understanding of variance and selection of principal component based on the variance. Also, visualization of PCA using scree plot, biplot etc. would aid the interpretation process.
6. Finally, exporting the outcomes of PCA into a spreadsheet/excel (.xlsx) file for variable (parameter) selection will be done.

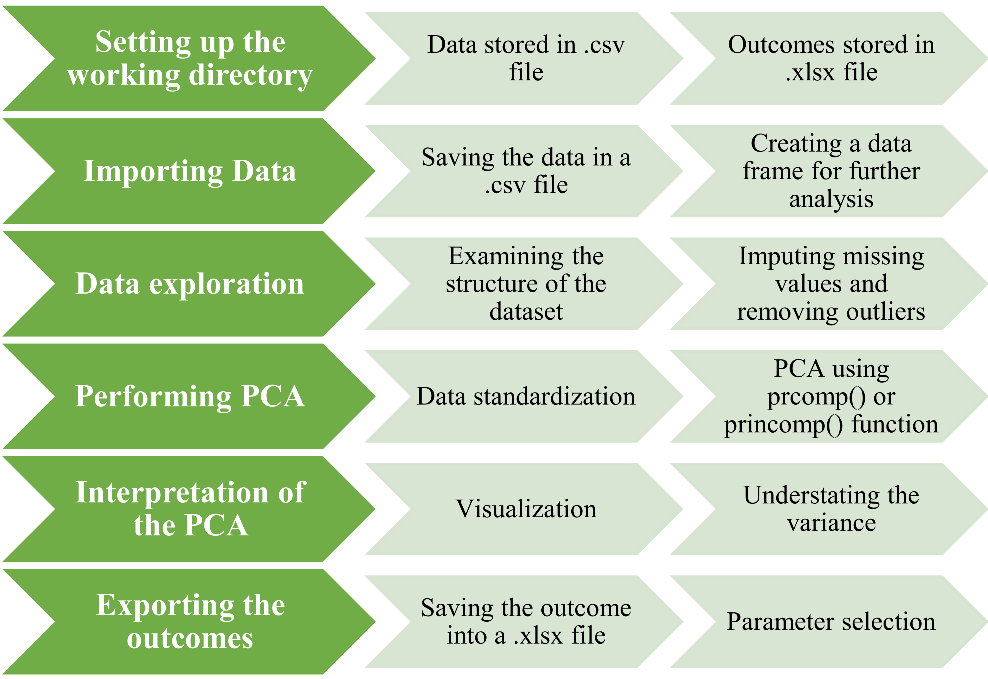


Figure 5.1: Steps involving in PCA using R

5.6 Development of the Water Quality Index Model

The WQI will be established to analyze natural and artificial activities based on fundamental groundwater chemistry markers. The water quality index can also assess the quality of river water in general as well as in terms of its intended use for drinking, recreation, aquatic zones, agriculture, etc. There are many water quality index calculation methods available for evaluating the water. In this study, the WQI model will be developed using the critical parameters obtained from PCA. The obtained WQIs from different models will be compared. Then, based on the research and PCA, a modified expression of WQI based on the NSF-WQI model will be proposed for Bangladeshi rivers, which will make the WQI calculations quick, objective, and reproducible and enable the evaluation of changes in water quality in various regions. To validate the performance of the proposed WQI model, the CCME model with the obtained critical parameters will be used.

5.7 Summary of Methodology

The research methodology has begun with sampling site selection. After ensuring the reliable sites for sample collection various water quality parameters will be analyzed. Based on the water quality parameter data, spatial interpolation and principal component analysis will be conducted. After determining the critical components from the principal component analysis, the WQI model will be developed based on the critical components and a modified expression for the WQI model based on the NFS-WQI model will be proposed for Dhaka city-based rivers which can be useful for further study of surface water quality of Bangladesh. The overall flow chart of the methodology is shown in as below –

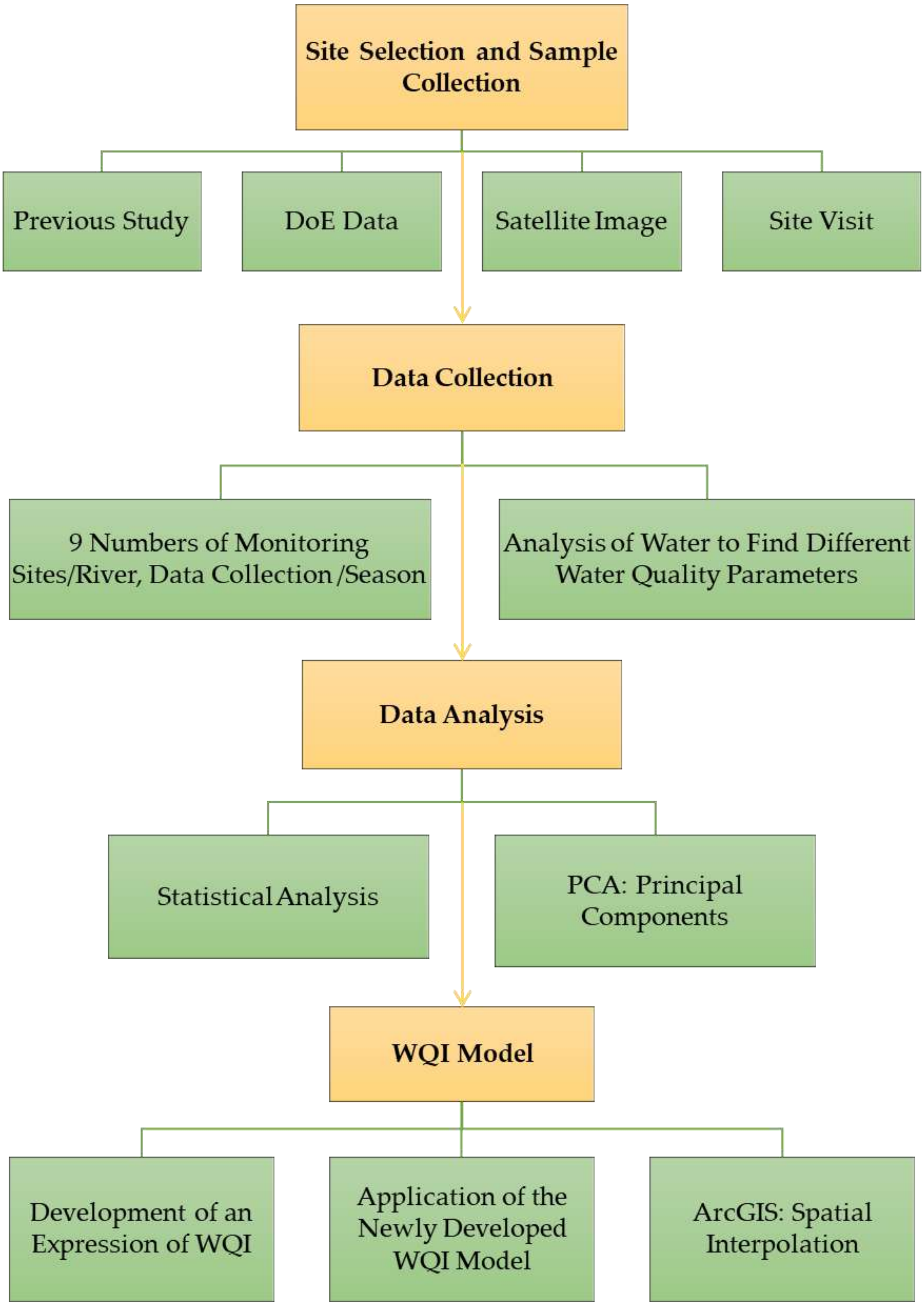


Figure 5.2: Flowchart for overall methodology

Chapter 6: Data Collection and Analysis

6.1 Collection and Analysis of Secondary Data: Case Study for 2019 Data

A secondary dataset of water quality parameters for four Dhaka-based rivers was collected from the Department of Environment (DoE) to establish a water quality index model with Principal Component Analysis (PCA) prior to water quality parameter selection and weighting. In the year 2019, 209 sets of measurements were done in 19 sampling stations during different months of the year by DoE, Bangladesh. The dataset covers a total of 12 WQ parameters such as temperature (°C), pH, electric conductivity (µS/cm), total alkalinity (mg/L as CaCO3), chloride (mg/L), total solids (mg/L), total dissolved solids (mg/L), total suspended solids (mg/L), dissolved oxygen (mg/L), biochemical oxygen demand (mg/L), chemical oxygen demand (mg/L) and turbidity (NTU) for the rivers Balu(Bl), Buriganga(Bg), Shitalakhya(Sl) and Turag(Tr).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Month	RiverNam	StationLal	Temp	pH	EC	TA	Cl	TS	TDS	TSS	DO	BOD	COD	Turb
2	January	Buriganga	Bg-1	21.7	7.52	661	220	25	381	345	36	0.76	12	40	23
3	January	Buriganga	Bg-2	22.1	7.27	714	224	24	393	371	22	0.81	18	45	11
4	January	Buriganga	Bg-3	22.1	7.44	712	226	23	406	376	30	0.96	14	38	12
5	January	Buriganga	Bg-4	22.3	7.41	701	232	24	386	370	16	0.7	12	32	14
6	January	Buriganga	Bg-5	22.5	7.41	699	244	22	407	377	30	0.46	14	48	10
7	January	Buriganga	Bg-6	22.5	7.65	690	242	23	396	374	22	0.23	18	50	12
8	January	Buriganga	Bg-7	22.4	7.69	710	240	24	405	344	61	1.2	16	42	33
9	January	Turag	Tr-1	21.9	7.98	333	270	10	190	170	20	0.17	25	72	15.4
10	January	Turag	Tr-2	23.5	7.64	331	274	11	209	163	46	0.31	27	97	30.6
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
201	December	Turag	Tr-2	18	7.34	469	183	30	347	264	83	0	12.4	66	59
202	December	Turag	Tr-3	17.6	7.73	679	192	46	440	387	53	0	12.2	56	50
203	December	Turag	Tr-4	18	7.69	617	250	56	442	350	92	0	8.6	39	70
204	December	Shitalakhy	Sl-1	27.8	7.17	332	166	34	234	192	42	1	11	33	41.7
205	December	Shitalakhy	Sl-2	21	7.32	230	120	18	162	154	18	4.5	6.2	16	17
206	December	Shitalakhy	Sl-3	21.3	7.34	330	162	32	210	175	35	2	8	30	35.3
207	December	Balu	Bl-1	19.1	8.37	931	312	57.5	593	515	78	0	16	72	43.9
208	December	Balu	Bl-2	17.7	7.44	1045	255	123	691	602	89	0	12	60	65
209	December	Balu	Bl-3	17.8	7.19	339	166	33.5	378	199	179	0.8	10.2	52	161
210	December	Balu	Bl-4	16.3	7.19	576	136	63	441	339	102	1.4	8.6	42	73

Figure 6.1: Collected secondary data from DoE

After collecting the data, Kaiser-Mayer-Olkin (KMO) test and Bartlett's test were done to check if the dataset is suitable for PCA analysis. Measure of Sampling Adequacy (MSA) value should be greater than 0.5 for KMO test and Bartlett's test shows the correlation exists in the dataset. Both the test results suggested that the dataset is suitable for PCA.

```
> KMO_result

── Kaiser-Meyer-Olkin criterion (KMO) ──────────────────────

✓ The overall KMO value for your data is meritorious.
  These data are probably suitable for factor analysis.

  Overall: 0.853

  For each variable:
 Temp    pH    EC    TA    Cl    TS   TDS   TSS    DO   BOD
0.150 0.826 0.903 0.839 0.906 0.835 0.804 0.815 0.899 0.882
  COD  Turb
0.874 0.951
> Bartletts_Result

✓ The Bartlett's test of sphericity was significant at an alpha
level of .05.
  These data are probably suitable for factor analysis.

  χ²(66) = 2687.89, p < .001
```

Figure 6.2: KMO and Bartlett's test results

The dataset was then standardized and scaled as the parameters have different units and different value ranges. PCA was performed on the standardized dataset using R programming language as R is specially designed for statistical analysis.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1		PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12
2	PC Variance Understanding												
3	Eigenvalu	6.680686	1.116108	1.084419	0.948462	0.574814	0.44453	0.433242	0.314726	0.233011	0.13623	0.025456	0.008318
4	Standard	2.584702	1.05646	1.041354	0.97389	0.758165	0.666731	0.658211	0.561004	0.482712	0.369094	0.159548	0.091203
5	Proportio	0.55672	0.09301	0.09037	0.07904	0.0479	0.03704	0.0361	0.02623	0.01942	0.01135	0.00212	0.00069
6	Cumulati	0.55672	0.64973	0.7401	0.81914	0.86704	0.90408	0.94019	0.96642	0.98583	0.99719	0.99931	1
7	Component loadings for PCs												
8	Temp	0.010422	-0.58704	0.671286	0.311617	-0.02421	-0.14929	0.091434	-0.18574	0.197487	-0.00561	0.047277	0.023329
9	pH	-0.21924	0.220189	0.592678	-0.27466	-0.32086	0.147423	-0.16542	0.447548	-0.34698	0.062048	0.012335	0.002521
10	EC	-0.36763	-0.07899	-0.06984	0.072659	-0.20949	0.018324	0.146437	-0.24376	-0.18149	0.142158	-0.78049	-0.24709
11	TA	-0.26367	0.514303	0.253438	-0.12927	-0.05368	0.074257	-0.13539	-0.43203	0.587571	-0.16232	-0.00056	-0.00795
12	Cl	-0.26295	-0.25046	-0.27812	0.360393	-0.4052	0.327804	-0.3735	0.30551	0.286211	-0.26981	0.040771	0.006749
13	TS	-0.36954	-0.06294	-0.11109	-0.02672	-0.1767	-0.05497	0.210167	-0.19815	-0.15965	0.083645	0.578174	-0.60165
14	TDS	-0.3658	-0.03684	-0.12118	0.02344	-0.22353	0.018006	0.169156	-0.2898	-0.21535	0.164826	0.213296	0.75344
15	TSS	-0.28556	-0.19093	-0.08472	-0.42506	0.026292	-0.22937	0.513449	0.423462	0.400182	-0.15905	-0.07478	0.08285
16	DO	0.243561	-0.32774	0.004978	-0.49278	-0.06932	0.711155	0.112046	-0.25207	0.014721	-0.05358	0.002231	-0.02602
17	BOD	-0.32717	0.00755	0.089178	0.121592	0.524863	0.185725	0.067356	-0.03336	-0.3331	-0.66534	-0.00951	0.028446
18	COD	-0.31851	-0.00336	0.03841	0.156826	0.529297	0.382273	-0.01597	0.200252	0.189152	0.607235	0.033734	-0.01333
19	Turb	-0.24141	-0.35238	-0.08942	-0.45982	0.207828	-0.31678	-0.65669	-0.13591	-0.03942	0.054923	-0.01354	0.010422

Figure 6.3: The outcome from principal component analysis

The scree plot indicates the percent of total variance accounted from each principal component. Here, the first principal component (PC1) has a 55.7% of the total variance and the second and third principal component contributes 9.3% and 9.0% of the total variance, respectively. Thus, implying a sum of 74% of total variance within the first three principal components.

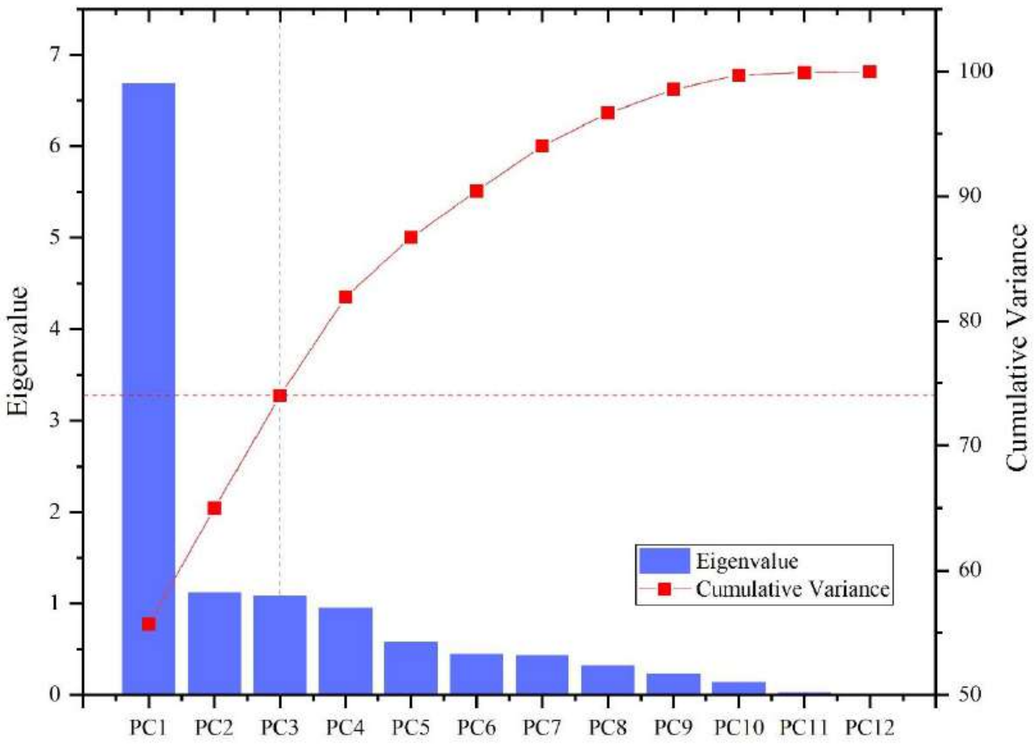


Figure 6.4: Scree plot as percent variance

The biplot represents the relations between the first two PCs. Each variable is represented with unique vectors indicating the contribution and their relation. For example, the horizontal axis represents the first PC, has a positive coefficient for temperature and dissolved oxygen whereas a negative coefficient for the other parameters. The second PC on the vertical axis has a positive coefficient for pH, total alkalinity and biochemical oxygen demand whereas a negative coefficient for the other 9 parameters.

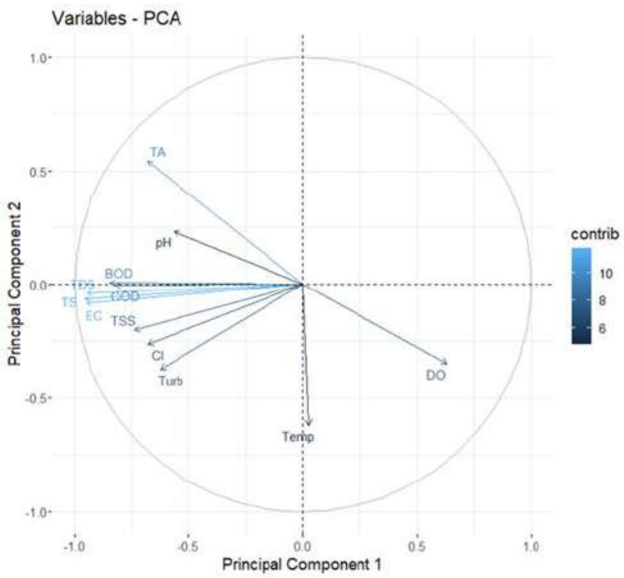


Figure 6.5: Biplot of first and second principal component

The variable contribution plot indicates the individual variable contribution for each principal component with a gradient color. Darker the color the more percent of contribution accounts for that variable. It was used for variable selection to reduce a higher degree of parameters.

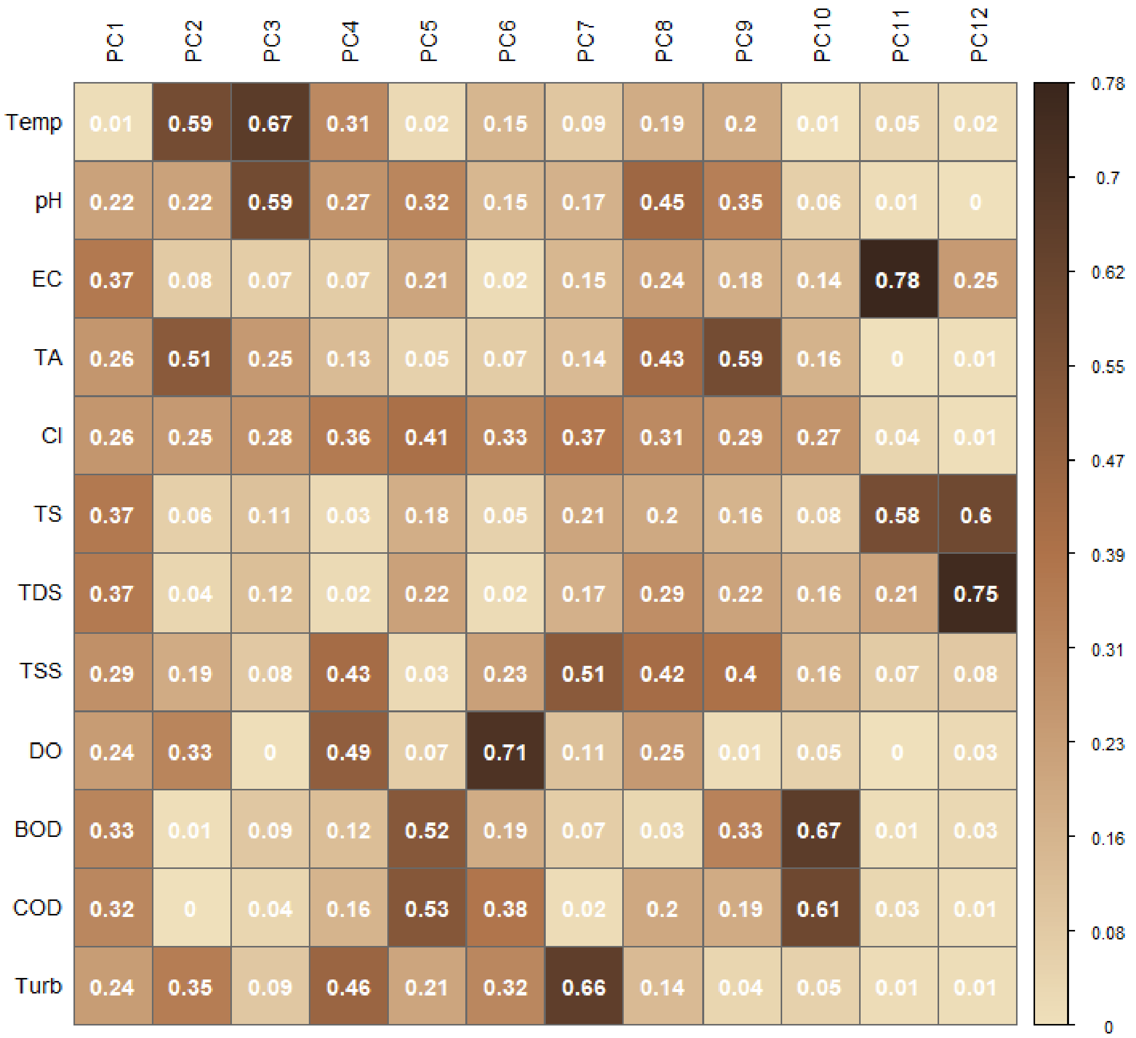


Figure 6.6: Absolute factor loading plot of PCA

For further parameter selection correlation analysis has been used. The correlation value > 0.8 is taken for discarding parameter. In this analysis TS, TDS, EC have a strong correlation value and TDS is selected from this group. BOD and COD have a higher correlation value. BOD is selected from this group.

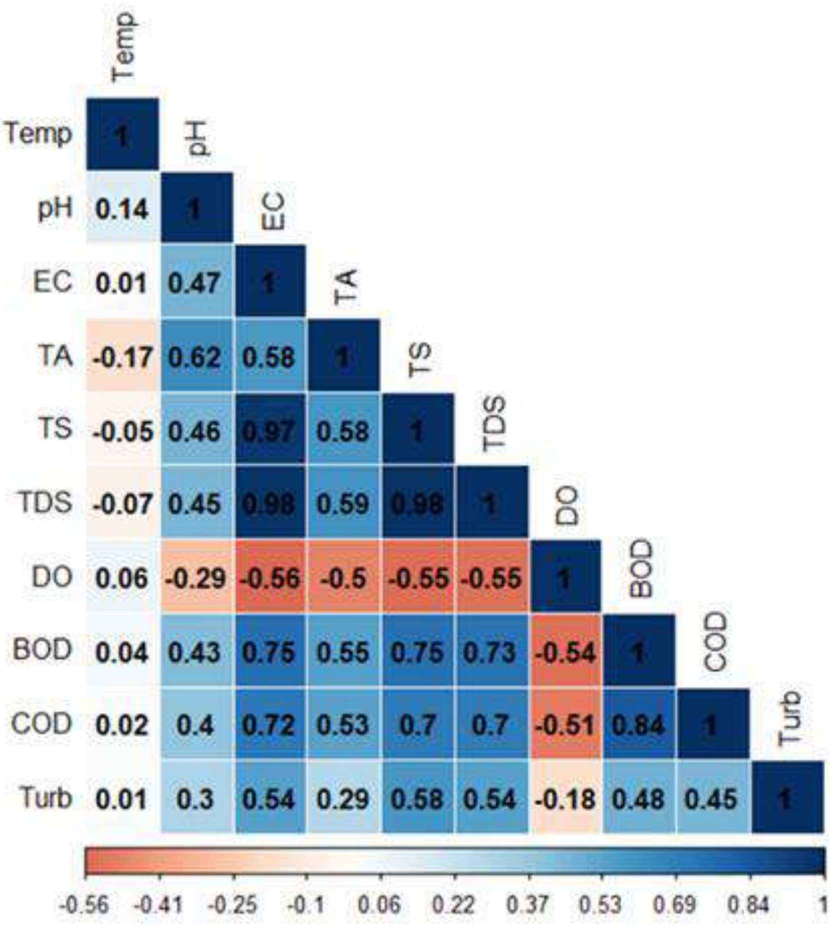


Figure 6.7: Correlation plot for the intermediate WQPs

PCA was performed again to calculate the parameter weights for the retained 7 original WQPs. Of the 7 PCs, 2 were selected that have the eigenvalue >1 accounting for 64.5% of the total variation. The loading of the selected parameters on the 2 PCs is shown in Table 6.1. The high absolute loading of pH, TA, TS, DO, BOD and turbidity were obtained in the first PC, and the absolute loading of temperature was higher in the second PC. The calculated weight for WQPs is presented in Table 6.2.

Table 6.1: Component loadings of the PCA for weight calculation

	Component Loading	
Eigenvalue	3.4245	1.0929
Relative Eigenvalue	0.7581	0.2419
Cumulative Variance (%)	48.92	64.54
Parameter/Variable	**PC1**	**PC2**
Temperature	0.019271	**0.907009**
pH	**-0.36713**	0.247783
TA	**-0.42901**	-0.18979
TS	**-0.47703**	-0.01616
DO	**0.368505**	0.204629
BOD	**-0.45762**	0.068231
Turbidity	**-0.32867**	0.182013

Table 6.2: Weight for 7 WQPs determined from PCA component loading matrix

PC	Relative Eigenvalue (E_r)	Parameter	Highest Absolute Loading	Relative Loading on Same PC (L_r)	Weight ($E_r \times L_r$)
1	0.7581	pH	0.3671	0.1512	0.12
		TA	0.4290	0.1767	0.13
		TS	0.4770	0.1965	0.15
		DO	0.3685	0.1518	0.12
		BOD	0.4576	0.1885	0.14
		Turbidity	0.3287	0.1354	0.10
		Total		1.0000	
2	0.2419	Temperature	0.9070	1.0000	0.24
		Total			1.00

Sub-indexing was done using the quality rating curves provided by the NSF-WQI model and modified Hallock (2002) method. The rating curves for each parameter are given below:

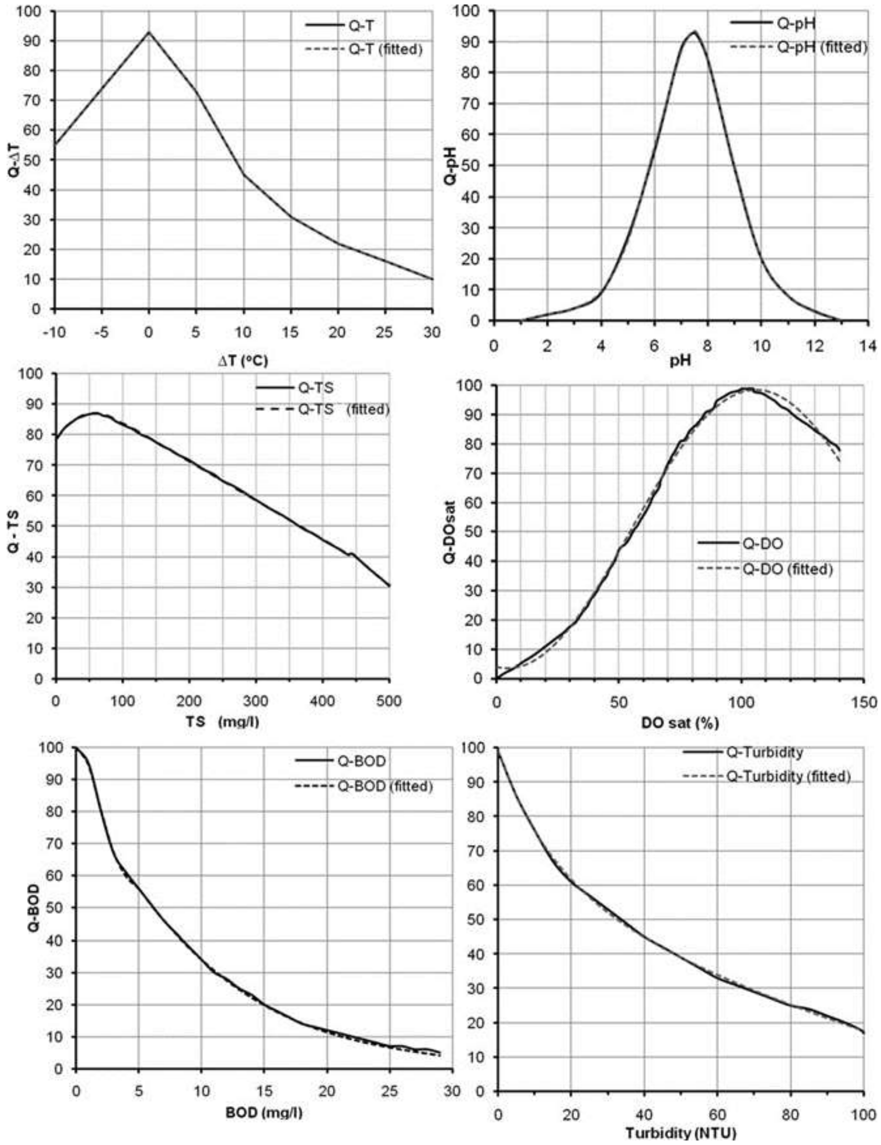


Figure 6.8: Rating curves from NSF-WQI

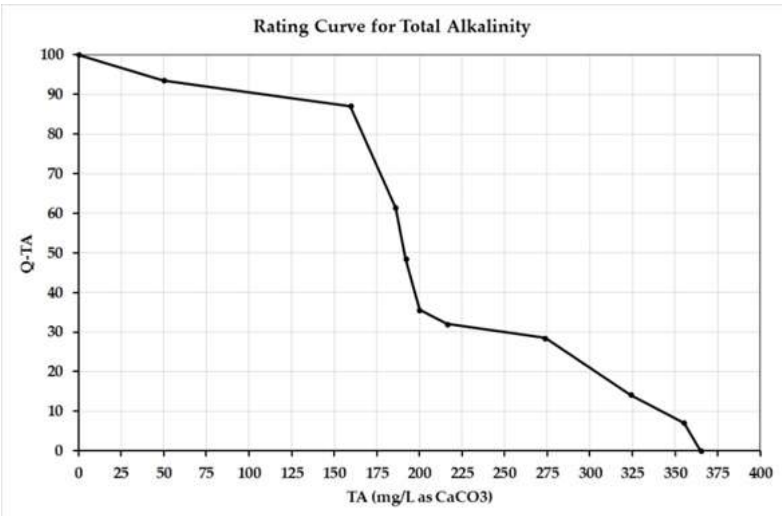
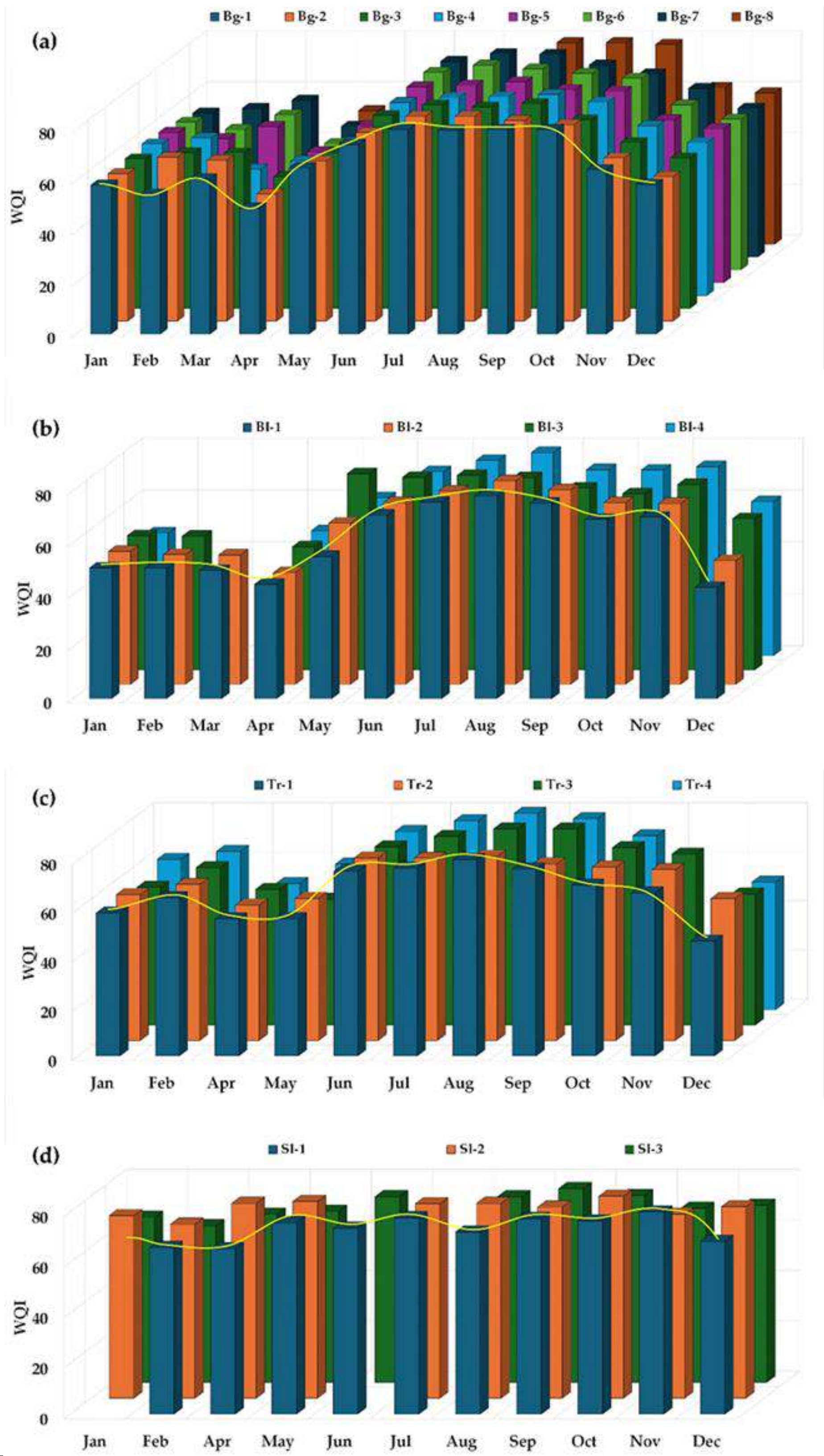


Figure 6.9: Modified rating curve for total alkalinity

The WQI values for each sampling station are shown below:



x

Figure 6.10: WQI of rivers (a) Buriganga (b) Balu (c) Shitalakhya and (d) Turag over the year 2019

The methodology can be best described as the following steps:

1. Data collection and parameter selection based on PCA and correlation analysis
2. Parameter weighting from the component loading data obtained from PCA
3. Sub-indexing using the quality rating curves provided by NSF-WQI model where applicable. If the rating curve does not exist for any specific parameter, then generate the rating curve using the method provided by Hallock with the data for that parameter.
4. Aggregating all the sub-indices into a single index which is called WQI and categorize using verbal and visual (color) remarks.

So, the overall methodology for calculating WQI can be expressed by the figure below:

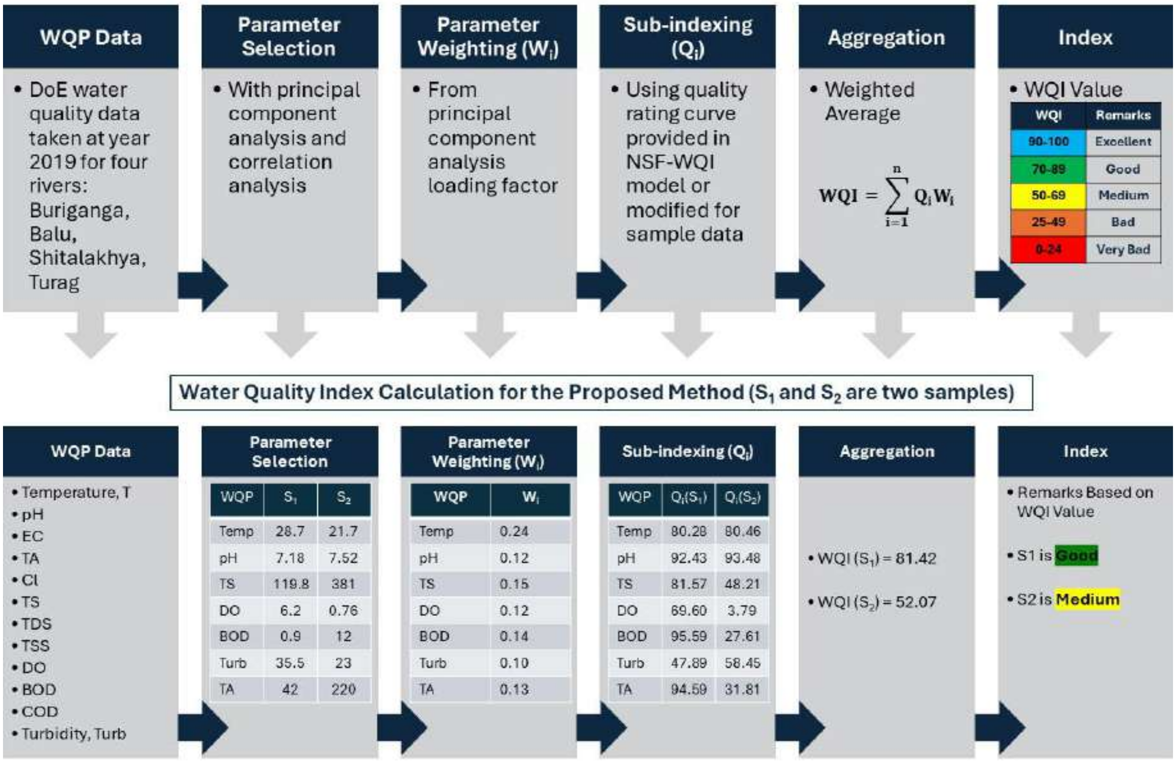


Figure 6.11: Overall methodology

The final outcomes of this case-study can be highlighted as followings.

1. A modified WQI model was developed with 7 reduced parameters from 12 WQPs for Dhaka-based rivers.
2. Correlation analysis followed by PCA can effectively reduce the number of parameters and ease weighting.
3. The WQI trend shows significant improvement in WQI value during wet season and the probable reason might be the heavy rainfall and dilution of water.
4. WQI for the rivers ranged from 42 to 84 over the year 2019 found from the model.
5. A WQI model can be constructively used in decision-making by the legislative authority of Bangladesh.

6.2 Site Selection

Site selection is mainly based on the Department of Environment's (DoE) sample collection points as well as on-site field visits. The Department of Environment has established a comprehensive monitoring network to diligently track surface water quality and this network entails collecting monthly samples from specific sampling stations strategically placed along rivers that are part of the network. Site selection is done for all four rivers. Initially 6 sampling sites were selected for each river. The sampling sites were increased from 6 to 9 for each river as per suggestions from the technical committee meeting.

The selected sampling locations are showed in the Table 6.3 with the reasons behind the selection process and Figure 6.12 shows all the selected sampling locations.

Table 6.3: Site selection reason

Site No.	River Name	Site Code	Sampling Sites	Reasons Behind Selection
1	Buriganga	BG1	*Gabtoli*	Nearby brick storage, steamer ghat
2		BG2	*Basila*	Nearby newly developed congested residential area
3		BG3	*Gudaraghat, Hazaribag*	Nearby former tannery industries
4		BG4	*Kholamora Ghat, Kamranghirchar*	Municipal discharge and transportation, recently sucker fish have been found and are in visible amounts on this site
5		BG5	*Chandir Ghat*	Nearby plastic recyclers, secondary recycling shops and factories
6		BG6	*Sadarghat Launch Terminal*	One of the busiest and most important areas in Buriganga
7		BG7	*Postogola Bridge, Dhaka*	Nearby small and large iron recycling shops and factories and an unknown liquid discharge line in the river
8		BG8	*Pagla Ghat, Dhaka*	Nearby brick resellers and food industries
9		BG9	*Fatullah Launch Terminal*	Launch terminal and nearby factories
10	Turag	TR1	*Kashimpur Namabazar*	Nearby public bazar and textile factories
11		TR2	*Mohisher Tek*	Nearby farms and agricultural field
12		TR3	*Ashulia Kacha Bazar*	Municipal waste disposal area beside the river
13		TR4	*Near Fulpukuria Thread, Tongi, Gazipur*	Nearby textile and chemical industries
14		TR5	*Tongi Rail Bridge, Gazipur*	Nearby public bazar, construction sites and hospitals
15		TR6	*Ashulia Landing Station*	The recent addition of the landing station and is the joining point of Turag and Tongi khal
16		TR7	*Rustompur Ghat, Mirpur road*	Household waste discharge
17		TR8	*Birulia*	Busy river crossing route
18		TR9	*Diabari Ghat, Mirpur road*	Nearby public bazar and the discharged municipal liquid waste into the river

19	Balu	BL1	*Pubail Bridge*	Nearby rural residential area
20		BL2	*Ulukhola Bazar*	Nearby public bazar and newly bridge construction site
21		BL3	*Talia Termukh Bridge*	Agricultural land, joining point of Balu and Turag
22		BL4	*Near Balu Bridge, 300 feet, Dhaka*	Nearby bazar and future cricket stadium, a prime attraction for tourists
23		BL5	*Near LGED Bridge, Jolshiri*	Nearby recreational spot- jolshiri and industrial zone Pran-RFL city
24		BL6	*Beraid Boat Ghat, Beraid*	The recent addition of the nearby Balu bridge and public bazar
25		BL7	*Near Edarkandi-Fakirkhali Road*	Nearby hybrid poultry farms by local people
26		BL8	*Near Eastern Straw Board and Paper Mill*	Nearby paper mills and residential area
27		BL9	*Rajakhali Ghat, Demra*	Nearby demra where Shitalakhya and Balu meets together, there is a nearby residential area and public bazar
28	Sitalakhya	SL1	*Kholapara-Ghorashal Kheyaghat, Ghorashal*	Nearby several painting and chemical industries and industrial discharge
29		SL2	*Kaliganj Kazir Char Kheyar Ghat, Kaliganj*	Nearby newly developing AK Khan economic zone
30		SL3	*Beldi Bazar, Beldi*	Nearby brick manufacturing industry
31		SL4	*Near Hatabo Bazar, Rupganj*	Nearby paper mills
32		SL5	*Habib Ghat, Boralu Bazar*	Industrial zone, several paper mills, food industries
33		SL6	*South Rupshi Masjid Ghat, Rupshi*	Nearby public bazar with small waste dumps
34		SL7	*Haripur Power Plant*	Nearby power plants and industrial zone
35		SL8	*Nabiganj Ghat*	Feri ghat and public bazar
36		SL9	*Khal Ghat*	Nearby wholesale bazar and steamer ghat

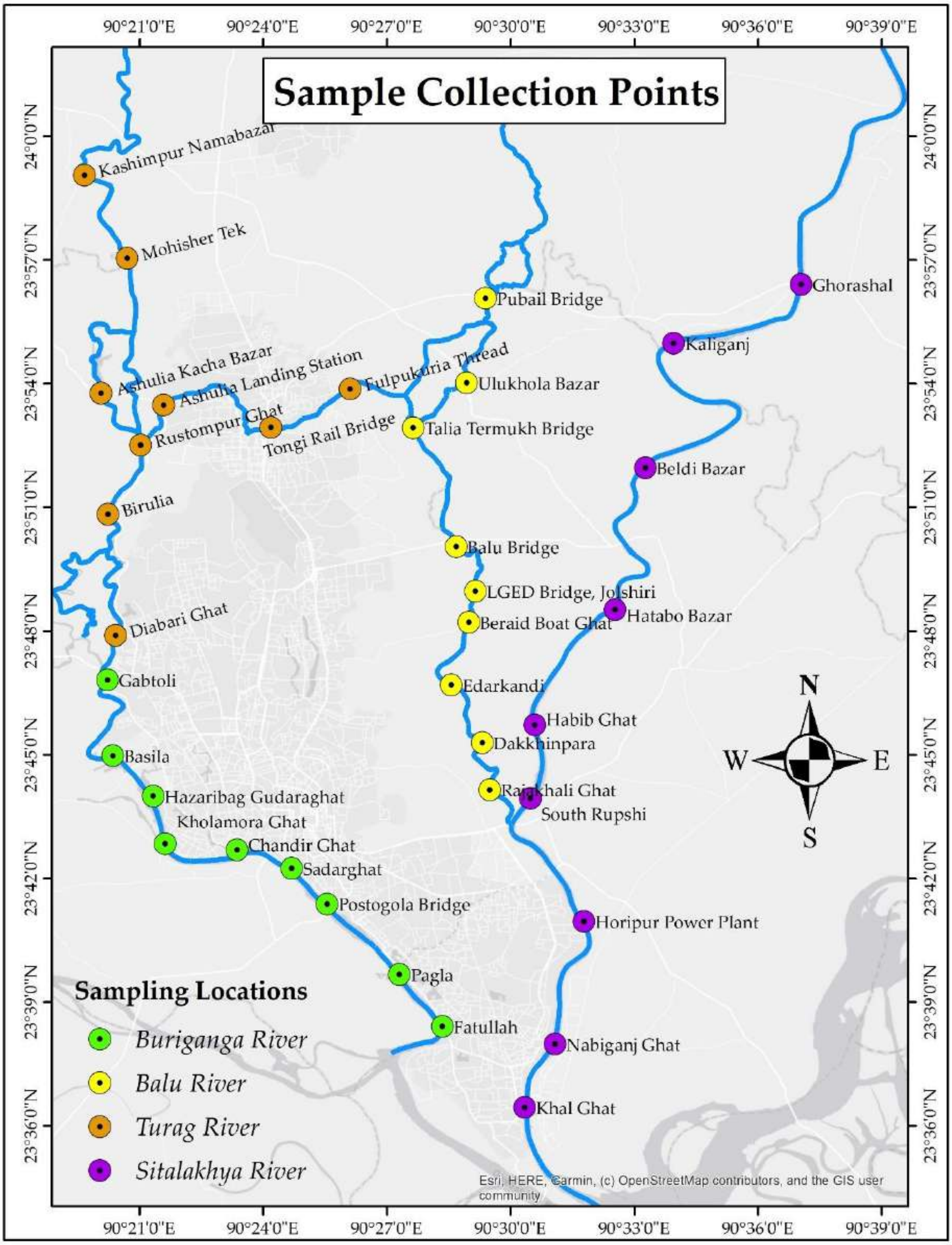


Figure 6.12: Sampling site locations for rivers

6.3 Preliminary Sample Collection and Analysis

Primary sample collection was intended to start from July, 2023 with collecting the monsoon data. However, due to the delay in equipment purchasing full sample collection had to start from January, 2024 by collecting the data for dry season. Meanwhile, before the purchase of equipment several sample collection and field visits were done during the monsoon and post-monsoon season and partial data were collected. To be specific, data were collected against single collection point for each river with limited lab facilities. Following is the partial dataset from the samples collected in monsoon and post-monsoon.

Table 6.4: Partial dataset from samples collected during monsoon and post-monsoon

Season	Station	Temp	pH	EC	TDS	DO	Turb	TSS	TS	BOD	COD	Cl
Moonsoon	BG5	26.7	6.9	192	91.2	2.2	10.5	19	110.2	5.5	14	16
Moonsoon	TR4	28.7	7.12	176	89.1	4.8	37.9	93	182.1	2.7	11	23
Moonsoon	BL4	29.5	7.29	189	89.4	4.3	35.8	40	129.4	5.1	24	28
Moonsoon	SL1	23.2	7.38	163	82.7	4.7	20.1	24	106.7	4.9	15	14
Post-Moonsoon	BG5	20.8	7.05	198	103	3.1	24.2	25	128	4.6	14	16
Post-Moonsoon	TR4	17.2	7.31	178	99.1	4.3	29.4	28	127.1	7	12	13
Post-Moonsoon	BL4	21.7	7.23	187	112.3	1.3	15.1	17	129.3	11	32	19
Post-Moonsoon	SL1	22.8	7.59	172	96.8	4.9	11.6	14	110.8	3.1	14	15

6.4 Primary Sample Collection and Analysis

After acquiring the necessary equipment, the research team successfully collected and analyzed samples from all designated sampling stations, resulting in a comprehensive dataset. Sample collection was carried out across four seasons in 2024. During the dry season, samples were gathered in January and February, with a total of six days dedicated to site visits for data collection. In the pre-monsoon season, sampling and onsite data collection occurred in May, with site visits spanning four days. For the monsoon season, sampling took place in July and August, with each river being visited for four days. Finally, in the post-monsoon season, samples were collected over five days in November.

The plots below present a detailed representation of how the parameter values vary across different seasons at each of the sampling sites.

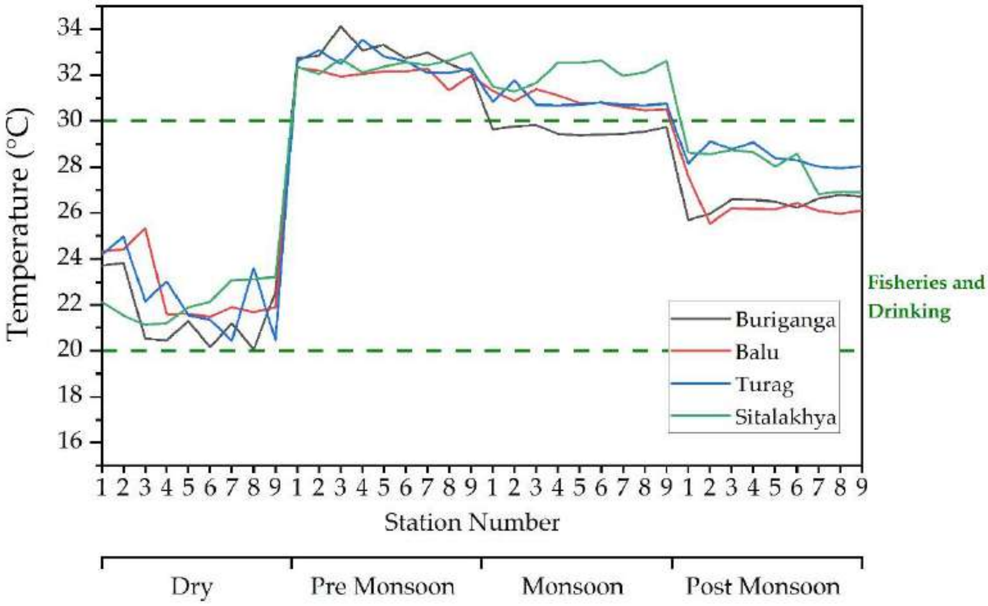


Figure 6.13: Variation of temperature during the four seasons

Figure 6.13 shows the variation in temperature across all four seasons. The plot indicates that during the dry season, the temperature was the lowest among all the seasons. In the pre-monsoon season, the temperature reached its highest point, and over time, it gradually decreased. This shows a clear trend where temperatures are higher in the pre-monsoon season, with a decreasing trend as the seasons progressed. The ideal temperature range for drinking and fisheries is in between 20 to 30 °C.

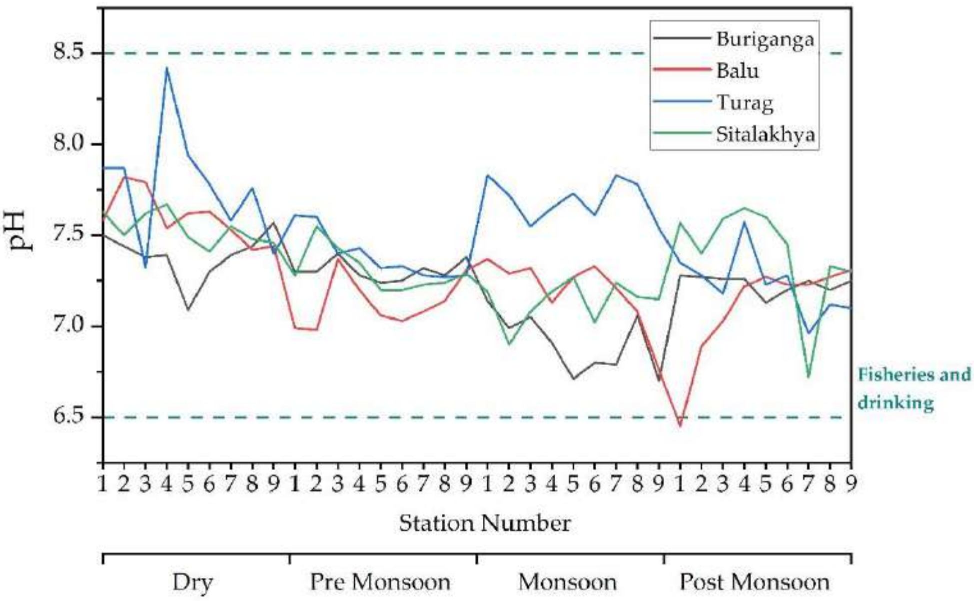


Figure 6.14: Variation of pH during the four seasons

Figure 6.14 illustrates the variation in pH across all four seasons. Throughout all seasons, the pH values consistently remained within the range of 7 to 8, with minor fluctuations observed at some sampling sites, showing either a slight increase or decrease in pH levels. The pH range for drinking water and fisheries is between 6.5 and 8.5.

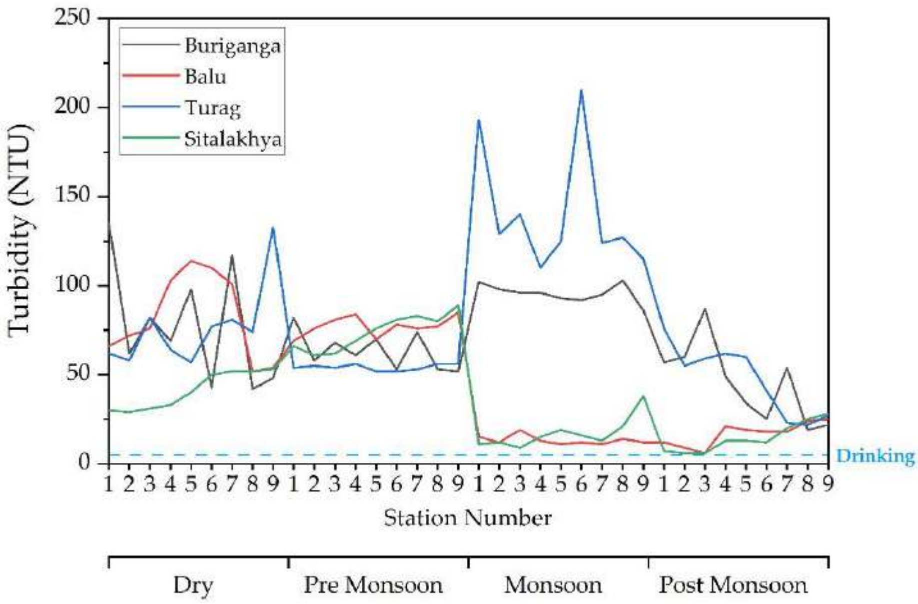


Figure 6.15: Variation of turbidity during the four seasons

Figure 6.15 illustrates the variation in turbidity across all four seasons. During the dry and pre-monsoon seasons, turbidity remained relatively stable, with only occasional peaks observed at some sampling stations. However, in the monsoon season, turbidity increased in the Buriganga and Turag rivers, while it decreased in the Balu and Sitalakhya rivers compared to the dry and pre-monsoon seasons. In the post-monsoon season, turbidity levels in the Balu and Sitalakhya rivers remained similar to those observed in previous seasons, whereas turbidity in the Buriganga and Turag rivers decreased.

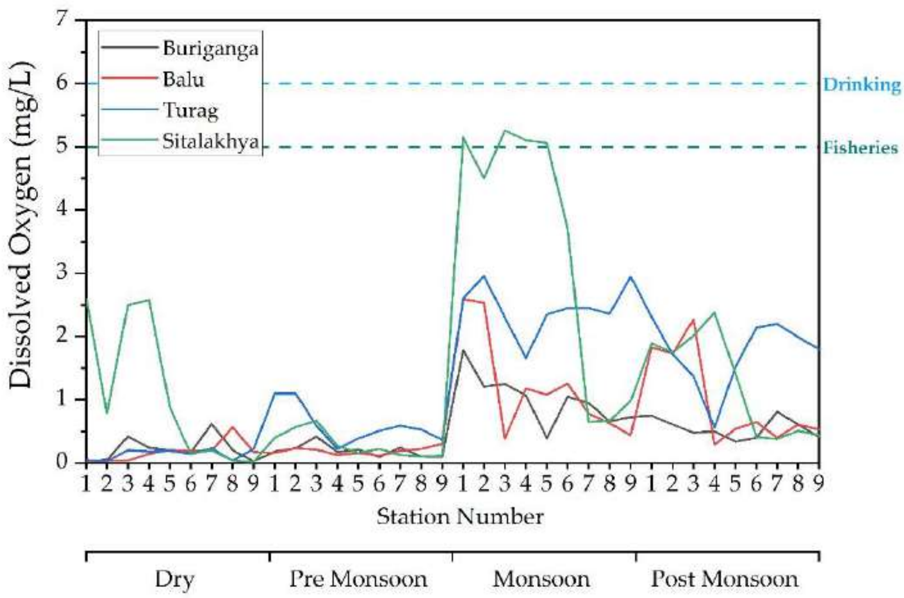


Figure 6.16: Variation of dissolved oxygen during the four seasons

Figure 6.16 illustrates the variation in dissolved oxygen (DO) across all four seasons. During the dry season, DO levels were at their lowest, representing the most critical condition. However, the upper stream of the Sitalakhya River exhibited higher DO levels compared to the other three rivers during this season. In the pre-monsoon season, a slight increase in DO was observed across all rivers. The monsoon season saw the highest DO levels in all rivers, with the most significant increase occurring in the Sitalakhya River. In the post-monsoon season, DO levels decreased but not to the extent seen during the dry season. For fisheries, the DO level should be greater than 5 mg/L, while for drinking water, it should exceed 6 mg/L. Nearly all sampling stations do not meet the required DO levels for fisheries, with the exception of the upper stream of the Sitalakhya River during the monsoon season.

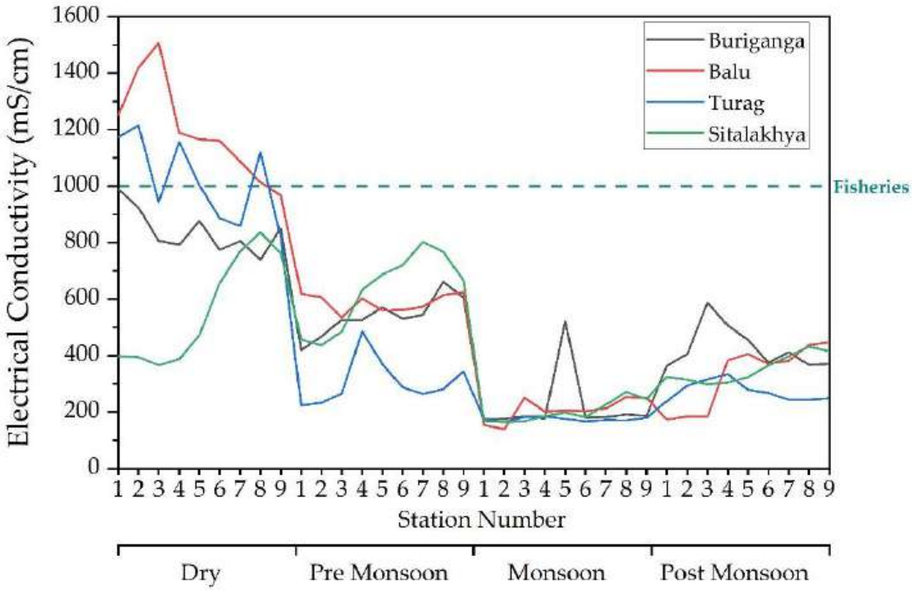


Figure 6.17: Variation of electrical conductivity during the four seasons

Figure 6.17 illustrates the variation in electrical conductivity (EC) across all four seasons. During the dry season, EC was highest in all rivers. Among them, the Balu River recorded the highest EC, while the Sitalakhya River exhibited the lowest. Notably, the Sitalakhya River showed an increasing trend in EC from the upper stream to the lower stream. Over time, EC levels decreased, reaching their lowest point during the monsoon season. In the post-monsoon season, EC levels began to increase again across all rivers. The EC for fisheries should generally be below 1,000 µS/cm, although this can vary depending on the type of fish being cultivated. With the exception of the Balu and Turag rivers during the dry season, all other sampling stations maintain an EC level below this threshold.

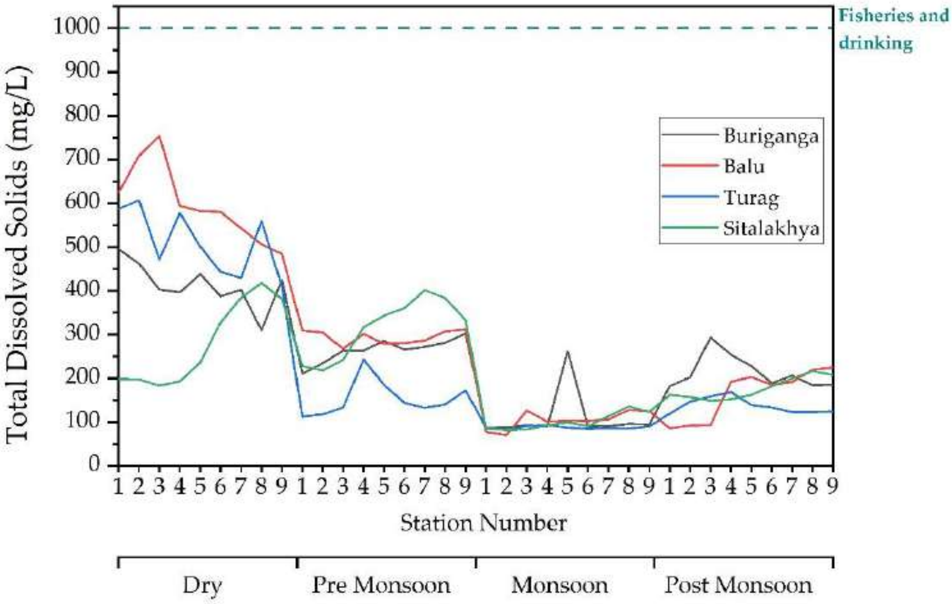


Figure 6.18: Variation of total dissolved solids during the four seasons

Figure 6.18 shows the variation in total dissolved solids (TDS) across all four seasons. TDS exhibited similar patterns to those observed in electrical conductivity (EC), following comparable trends throughout the seasons. The standard for TDS for both drinking water and fisheries is 1,000 mg/L. All the sampling stations are within this limit.

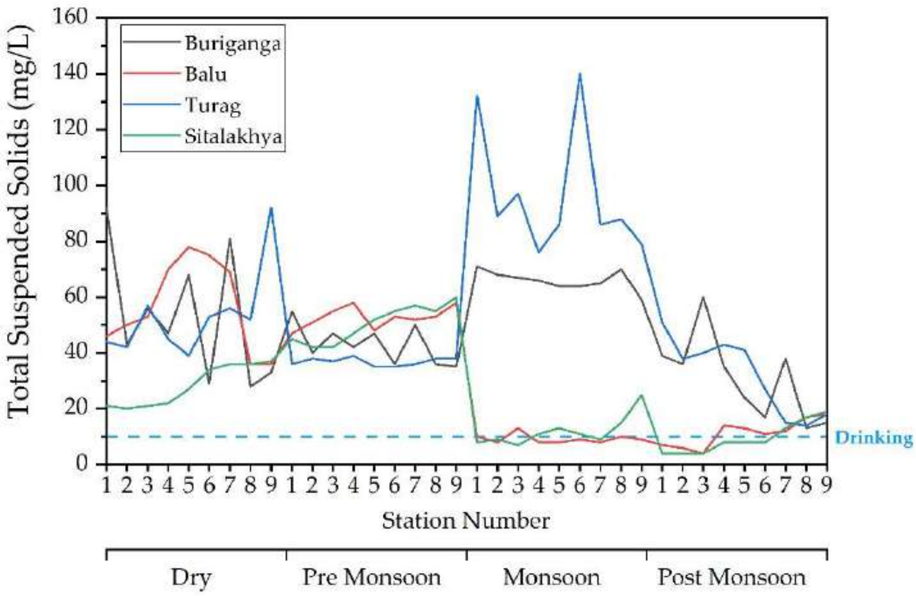


Figure 6.19: Variation of total suspended solids during the four seasons

Figure 6.19 shows the variation in total suspended solids (TSS) across all four seasons. During the monsoon season, the Turag River recorded the highest levels of suspended solids throughout the year. In contrast, during the dry and pre-monsoon seasons, TSS levels remained relatively stable, with only occasional spikes at certain sampling sites. In the monsoon season, TSS levels decreased in the Balu and Sitalakhya rivers, while in the Buriganga River, TSS increased, although not to the same extent as in the Turag River. The limit for TSS in drinking water should be below 10 mg/L. During the monsoon season, both the Balu and Sitalakhya rivers were observed to be around this range.

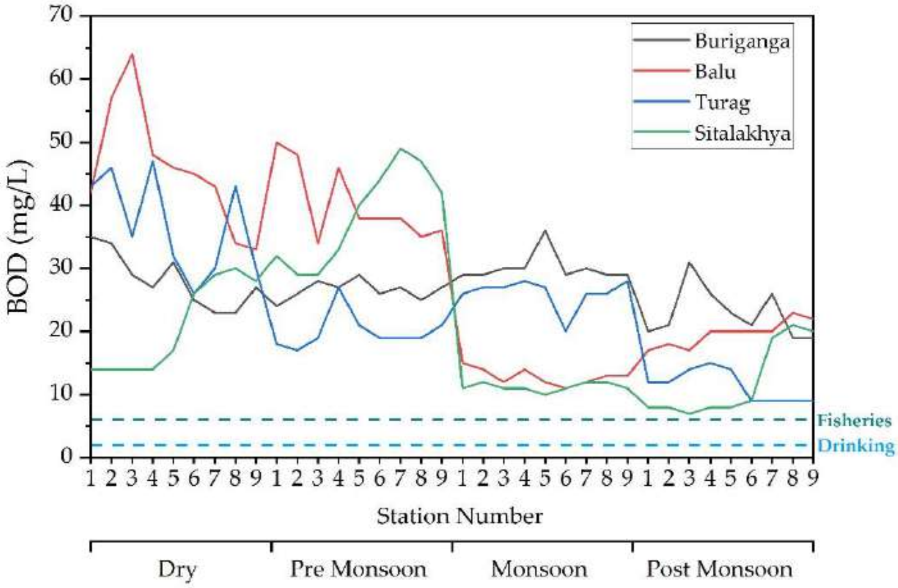


Figure 6.20: Variation of biological oxygen demand during the four seasons

Figure 6.20 illustrates the variation in biological oxygen demand (BOD) across all four seasons. During the dry and pre-monsoon seasons, BOD values were highest in all rivers, except for the Sitalakhya River. The Balu River recorded the highest BOD during the dry season, with a slight decrease observed in the pre-monsoon season. In contrast, the Sitalakhya River showed an increasing trend in BOD from the dry season to the pre-monsoon season. During the monsoon, BOD levels decreased in both the Balu and Sitalakhya rivers, while the BOD remained unchanged in the Buriganga and Turag rivers. The limit for BOD is 2 mg/L for drinking water and 6 mg/L for fisheries. All the sampling stations exceed these BOD limits.

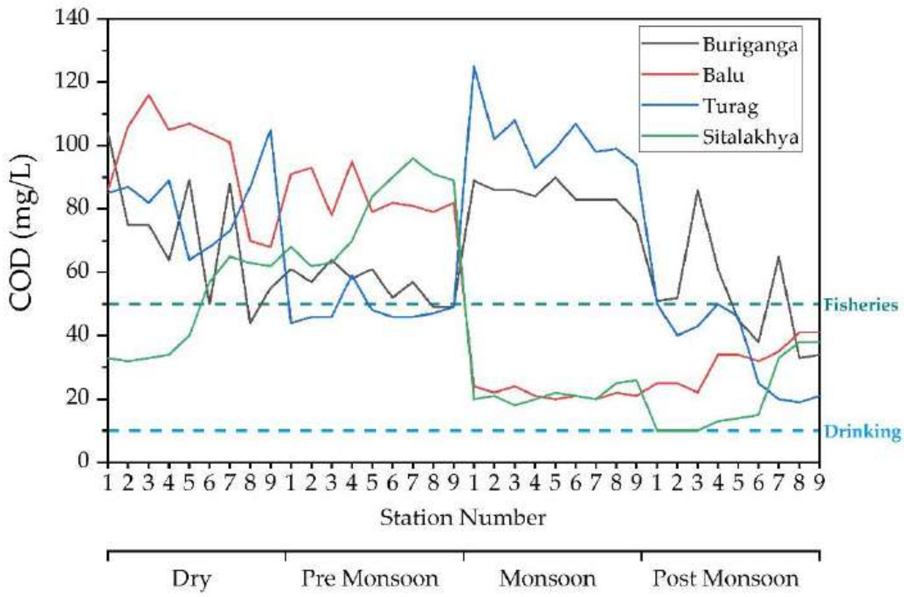


Figure 6.21: Variation of chemical oxygen demand during the four seasons

Figure 6.21 illustrates the variation in chemical oxygen demand (COD) across all four seasons. During the dry season, the Buriganga, Turag, and Balu rivers exhibited higher COD levels compared to the Sitalakhya River. In the Sitalakhya River, COD showed an increasing trend over time, reaching its peak during the pre-monsoon season. During the monsoon season, COD decreased in both the Balu and Sitalakhya rivers, while it increased in the Buriganga and Turag rivers. In the post-monsoon season, COD levels decreased in the Buriganga and Turag rivers compared to the monsoon season, whereas COD levels in the Balu and Sitalakhya rivers tended to increase. The threshold for COD is 10 mg/L for drinking water and 50 mg/L for fisheries. During the monsoon and post-monsoon seasons, both the Balu and Sitalakhya rivers have COD levels below the limit for fisheries. However, all sampling stations exceed the limit for drinking water.

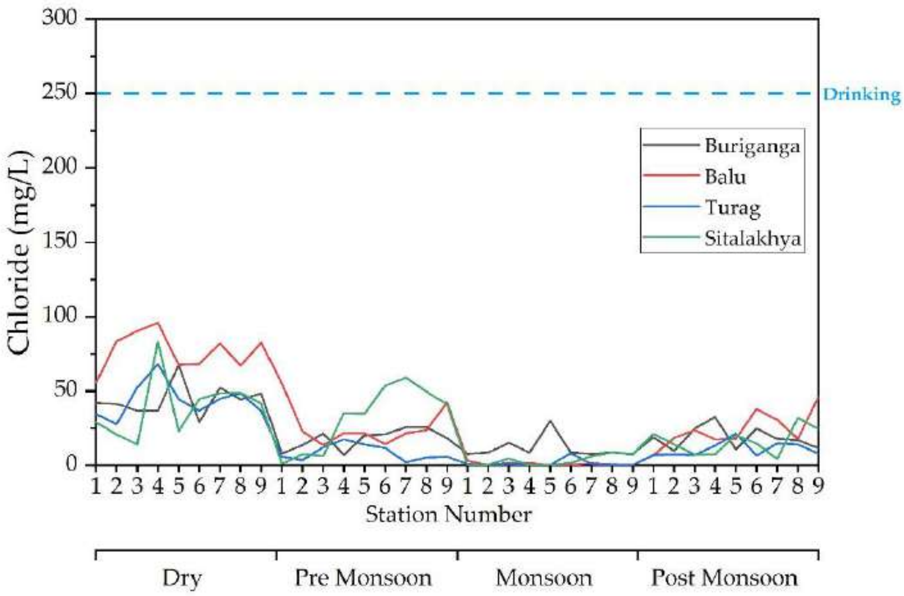


Figure 6.22: Variation of chloride ion concentration during the four seasons

Figure 6.22 illustrates the variation in chloride ion concentrations across all four seasons. During the dry season, all four rivers exhibited the highest concentrations of chloride ions at the sampling sites, with the Balu River showing the highest levels among them. In the pre-monsoon season, chloride ion concentrations decreased compared to the dry season, with the Sitalakhya River recording the highest chloride ion levels during this period. During the monsoon season, chloride ion concentrations decreased significantly in all rivers, except for the Buriganga River. In the post-monsoon season, chloride ion concentrations gradually increased in all rivers.

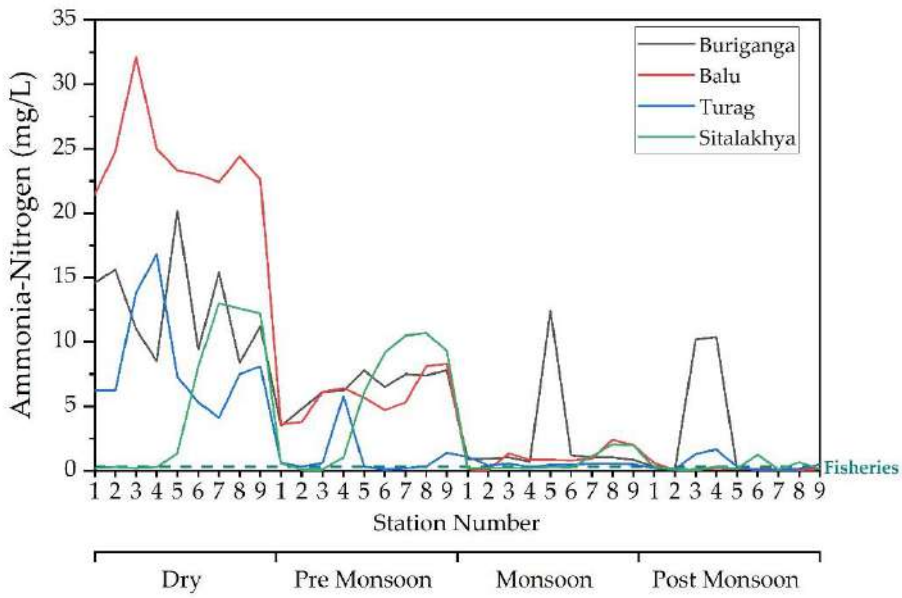


Figure 6.23: Variation of ammonia-nitrogen concentration during the four seasons

Figure 6.23 illustrates the variation in ammonia-nitrogen concentrations across all four seasons. The Balu River recorded the highest ammonia-nitrogen concentrations during the dry season, surpassing all other rivers and seasons. Over time, the concentration in the Balu River gradually decreased. In the Sitalakhya River, ammonia-nitrogen concentrations were higher in the downstream areas during the dry and pre-monsoon seasons, while the upper stream exhibited lower concentrations. During the monsoon and post-monsoon seasons, concentrations in the Sitalakhya River remained lower and showed minimal variation. In the Buriganga and Turag rivers, ammonia-nitrogen concentrations were highest during the dry season, with a gradual decline over time, although occasional spikes were observed at certain sampling sites in the Buriganga River.

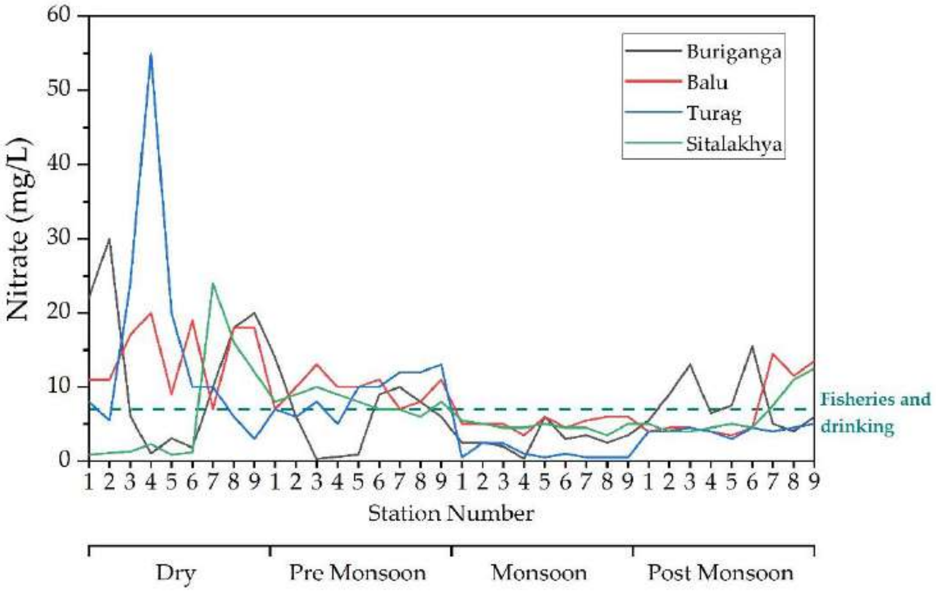


Figure 6.24: Variation of nitrate ion concentration during the four seasons

Figure 6.24 shows the variation in nitrate ion concentrations across all four seasons. The concentration of nitrate ions remained relatively stable throughout the year, with only minor fluctuations. During the dry season, the nitrate ion concentration was slightly higher compared to the other seasons, while it was the lowest during the monsoon. Overall, the concentrations remained consistent across the year, with the exception of a noticeable spike at one sampling station in the Turag River during the dry season. The threshold for nitrate is 7 mg/L for both drinking water and fisheries. During the monsoon season, none of the sampling stations across the four rivers exceeded this limit.

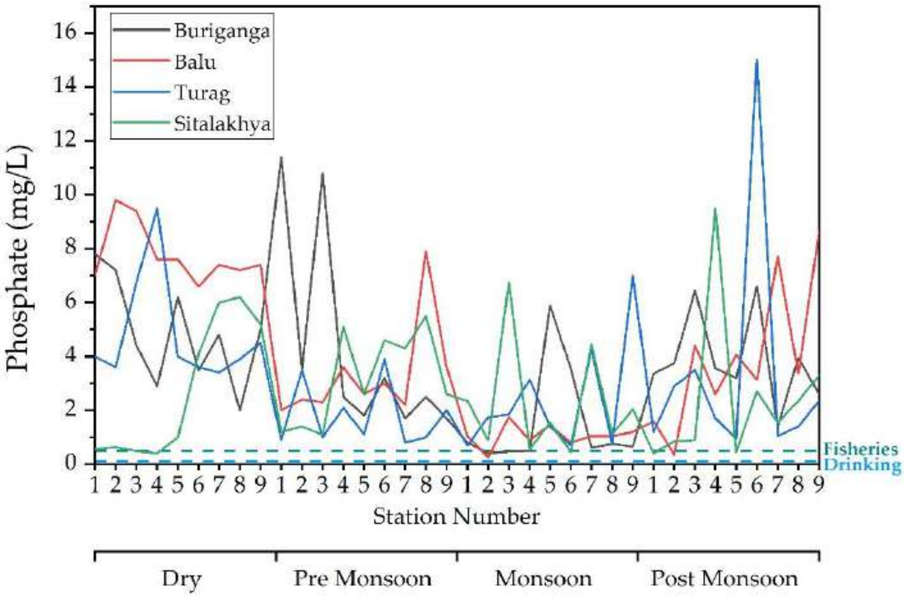


Figure 6.25: Variation of phosphate ion concentration during the four seasons

Figure 6.25 illustrates the variation in phosphate ion concentrations across all four seasons. The Balu River exhibited the highest phosphate ion concentration during the dry season. Throughout all seasons, the concentration of phosphate ions varied across different sampling stations. Notably, the highest concentration of phosphate ions was observed in the post-monsoon season at one sampling station in the Turag River. The limit for phosphate is 0.1 mg/L for drinking water and 0.5 mg/L for fisheries. All the sampling stations have phosphate levels higher than the threshold for fisheries.

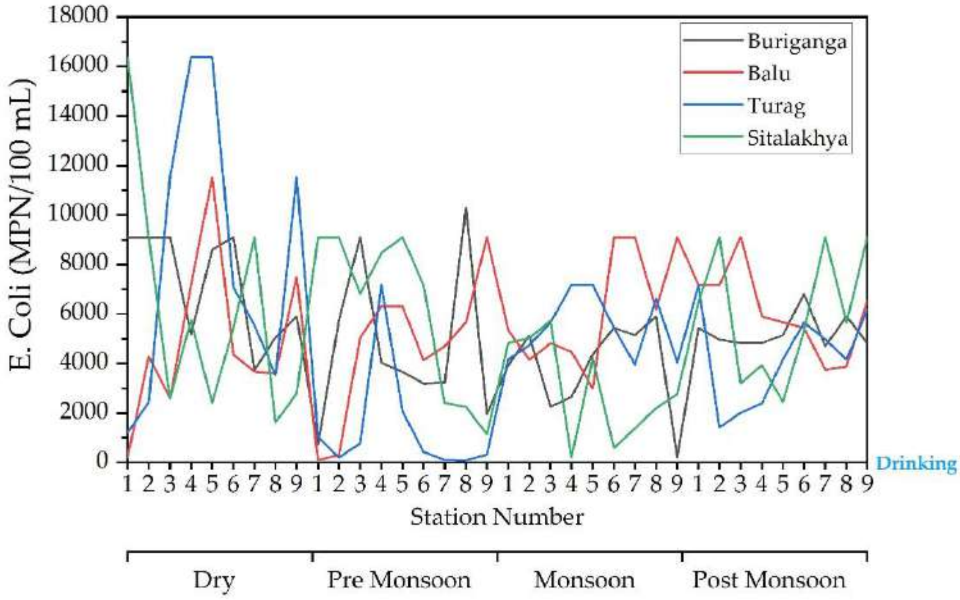


Figure 6.26: Variation of amount of E coli during the four seasons

Figure 6.26 shows the variation in the amount of E. coli across all four seasons. In all rivers, E. coli concentrations varied across the sampling stations in each season, without following any specific trends. For drinking E. coli should be 0 in the water.

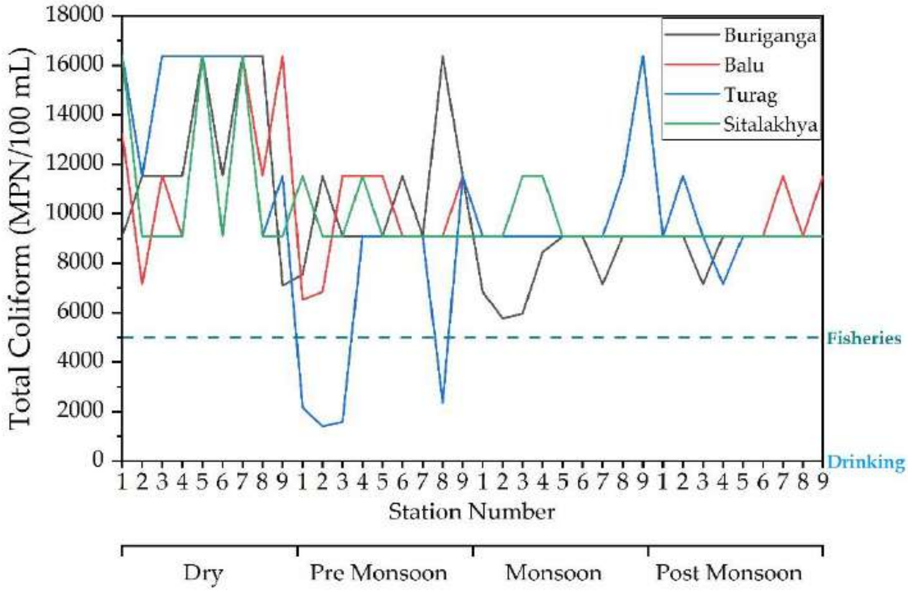


Figure 6.27: Variation of amount of total coliform during the four seasons

Figure 6.27 illustrates the variation in the amount of total coliform across all four seasons. During the dry season, total coliform levels were higher in all rivers. In the pre-monsoon season, these levels decreased. However, in the monsoon and post-monsoon seasons, total coliform concentrations remained largely unchanged, with the exception of occasional spikes. The total coliform count should not exceed 5,000 MPN/100mL for fisheries, and it should be 0 for drinking water. Almost all the sampling stations had total coliform levels higher than 5,000. However, in the pre-monsoon season, a few sampling stations along the Turag River exhibited total coliform counts below 5,000.

6.5 Statistical Analysis of the Collected Data

6.5.1 Box plot of water quality parameters

After collecting data from all sampling stations across the four rivers and four seasons, box and whisker plots were created for each water quality parameter, grouped by river, to visualize the distribution of the data. These plots are particularly useful for understanding the spread, central tendency, and variability of the parameters across the different rivers. To facilitate comparison between the parameters, the data were standardized into Z-scores, and box plots were regenerated using these standardized values.

Figure 6.28 shows the distribution of z-scores for water quality parameters across four rivers: Buriganga, Balu, Turag, and Sitalakhya, allowing for standardized comparison. The plots show notable differences in the variability and central tendencies of the parameters among the rivers. In many cases, the median is not centered within the interquartile range, and the mean deviates from the median, suggesting skewed distributions. Buriganga and Sitalakhya exhibit wider interquartile ranges for maximum parameters, indicating greater variability in the data. In contrast, Turag and Balu generally display comaparatively narrower interquartile ranges, suggesting more consistent parameter values. Buriganga, Turag, and Sitalakhya show a higher frequency of outliers, whereas Balu has fewer. Despite this, outliers are present across all rivers, with some parameters demonstrating a higher frequency of extreme values. These observations indicate significant differences in the overall distributions and variability of water quality parameters among the rivers.

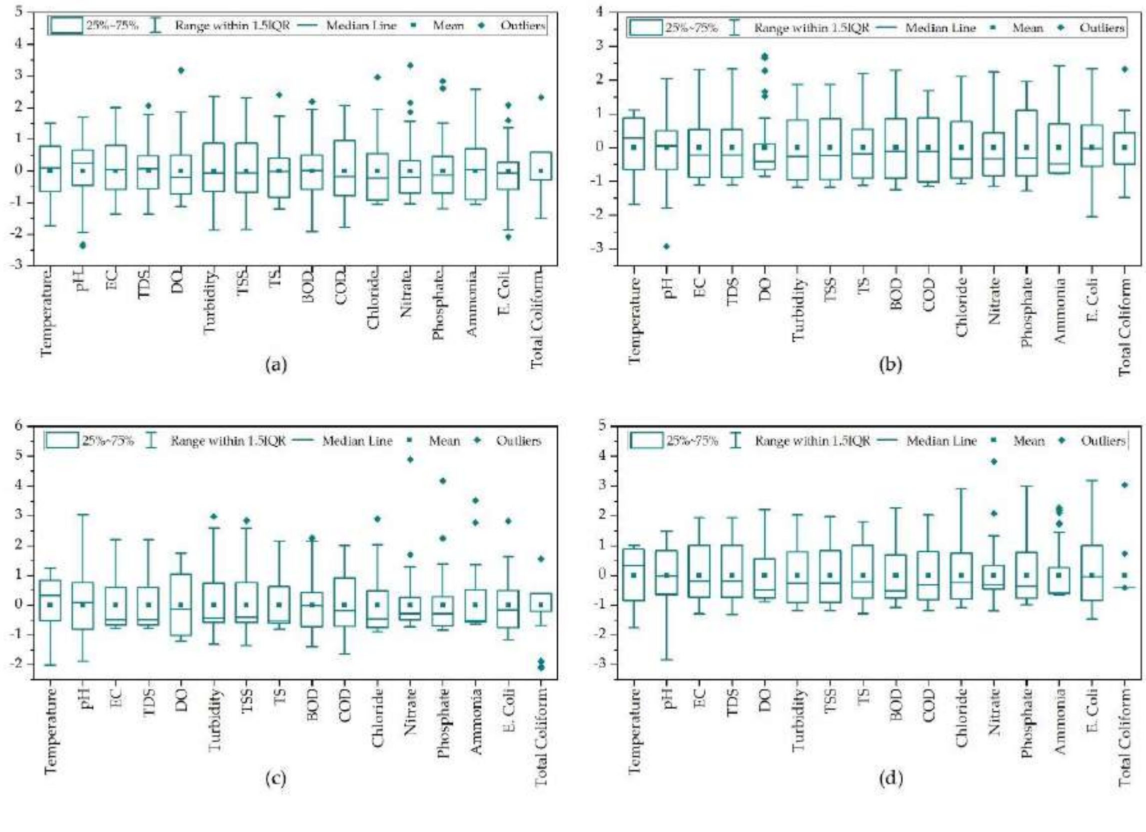


Figure 6.28: Z-value distribution of water quality parameters for (a) Buriganga (b) Balu (c) Turag, and (d) Sitalakhya River

6.5.2 Multivariate Analysis of Variance (MANOVA)

Multivariate Analysis of Variance (MANOVA) is a statistical technique used to examine the differences in multiple dependent variables simultaneously across different groups or levels of an independent variable. It is an extension of the ANOVA (Analysis of Variance) that allows researchers to assess how several outcomes are affected by one or more factors, while accounting for the intercorrelations among the dependent variables. MANOVA is commonly used in fields such as psychology, biology, and social sciences to understand complex relationships between multiple variables.

MANOVA was conducted with the collected data to examine the effects of Season and River on 16 water quality parameters, with data collected from four rivers, four seasons, and nine sampling stations per river. The analysis indicated that both independent variables significantly affected the multivariate response. The effect of Season was highly significant (Wilks' Lambda = 0.0040, $F = 44.07$, $p < 0.001$), suggesting considerable seasonal variation in water quality parameters. Similarly, the effect of River was also statistically significant (Wilks' Lambda = 0.1837, $F = 6.25$, $p < 0.001$), indicating spatial differences in water quality. The larger F-value and associated multivariate statistics for Season suggest that temporal variations exert a stronger influence on water quality patterns than spatial differences across rivers. The detailed result of the MANOVA analysis is given in the Table 6.5.

Table 6.5: Results found from MANOVA

Season	Value	Num DF	Den DF	F Value	Pr > F
Wilks' lambda	0.0040	45.000	366.1823	44.0703	0.0000
Pillai's trace	2.0496	45.0000	375.0000	17.9709	0.0000
Hotelling-Lawley trace	54.0015	45.0000	294.8551	146.1957	0.0000
Roy's greatest root	51.5703	15.0000	125.0000	429.7522	0.0000
River	**Value**	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Wilks' lambda	0.1837	45.0000	366.1823	6.2577	0.0000
Pillai's trace	1.1836	45.0000	375.0000	5.4305	0.0000
Hotelling-Lawley trace	2.6120	45.0000	294.8551	7.0712	0.0000
Roy's greatest root	1.6768	15.0000	125.0000	13.9735	0.0000

Chapter 7: Water Quality Index Modelling

7.1 Development of WQI model

The general structure of WQI models include four main steps, namely water quality parameter selection, generation of parameter sub-indices, parameter weighting, and final indexing using an aggregating function (Abbasi & Abbasi, 2012b), (Abrahão et al., 2010), (Lumb, Sharma, & Bibeault, 2011). Figure 7.1 illustrates the structure followed in this structure to develop the WQI model.

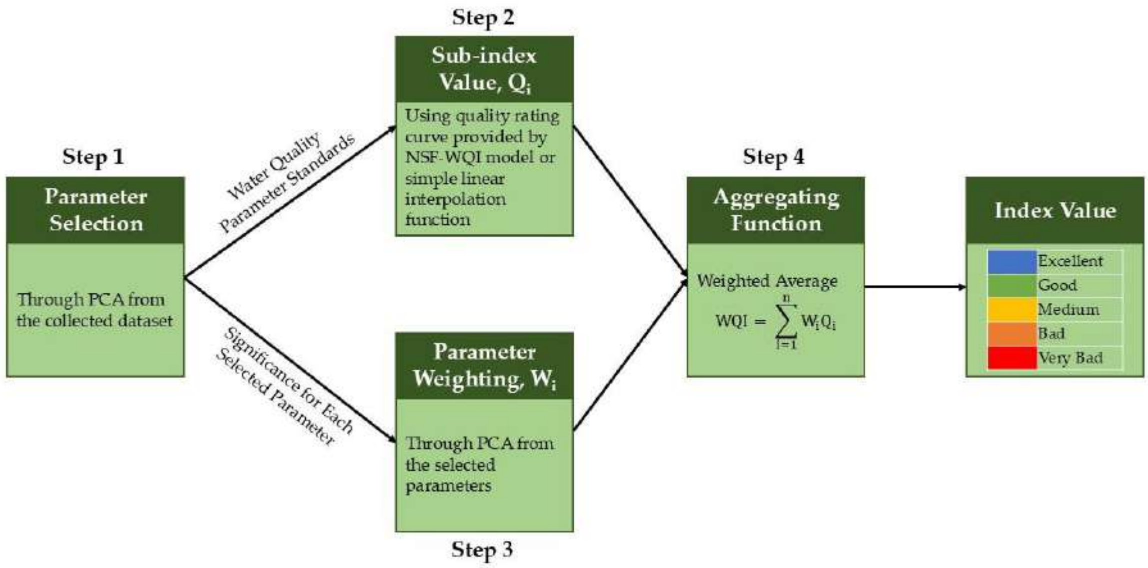


Figure 7.1: Structure for the development of WQI model

7.1.1 Parameter Selection

After collecting data from each sampling station of the rivers across the four different seasons, Principal Component Analysis (PCA) was performed. The dataset was standardized and scaled due to the different units and value ranges of the water quality parameters (WQPs). PCA was conducted using the R programming language, which is specifically designed for statistical analysis and includes built-in packages for performing PCA. In this study, R version 4.4.0 and RStudio 2024.04.0 were used for performing PCA.

From the PCA outcome, parameter selection was done based on the criteria below which was suggested by (Mandal et al., 2008), (Tripathi & Singal, 2019) and (Nath Roy et al., 2024).

i. The principal components (PCs) that explain the most variance was identified. This was done by identifying those PCs having eigenvalue greater than 1.
ii. Parameters with absolute loadings greater than 0.3 in those PCs were selected for this study, while the remaining parameters were discarded.

The correlation matrix of the selected parameters was examined. Parameters with a correlation value less than 0.9 were selected, while those with a correlation value greater than 0.9 were grouped. From each group, only one parameter was retained, and the others were discarded.

For the parameter selection, PCA was performed on the sixteen water quality parameters, and based on the eigenvalue criteria, the first three principal components (PCs) were selected. These three PCs collectively explain 74.06% of the variance. PC1 explains approximately 49% of the total variance, while PC2 and PC3 account for 16.41% and 8.69% of the variance, respectively. All three principal components have eigenvalues greater than 1, indicating their significance in capturing the variability within the data. Therefore, the first three PCs were selected to decrease the dimensionality. Figure 7.2 shows the scree plot of the PCA results, where the eigenvalues of each PC are represented by the bars, and the line indicates the cumulative variance explained by the PCs.

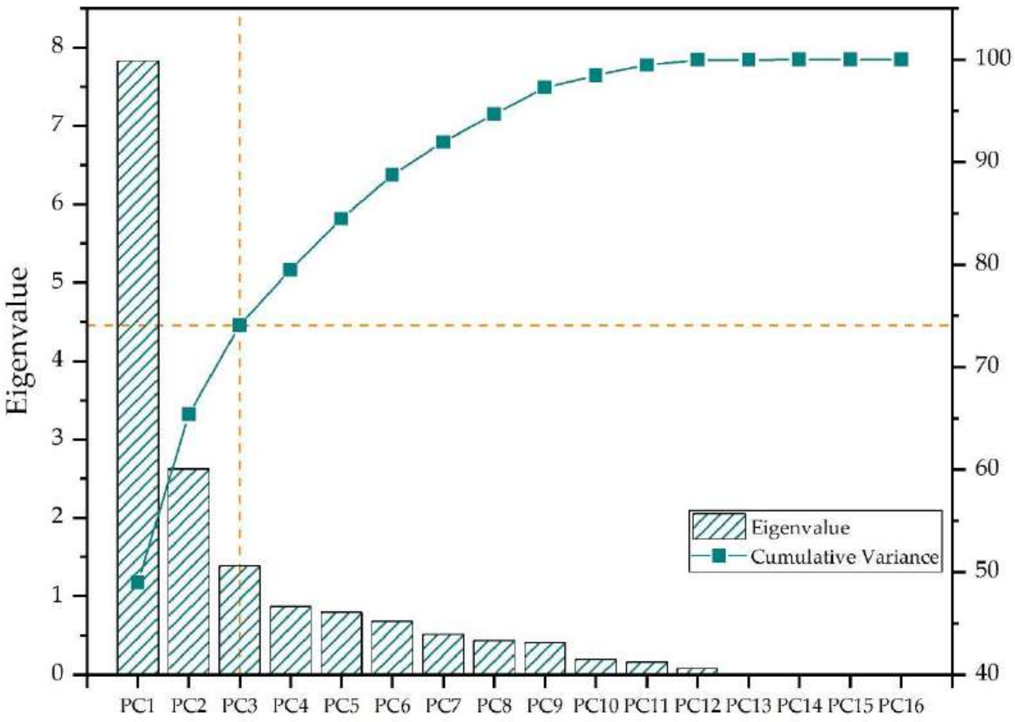


Figure 7.2: Scree plot showing eigenvalues and cumulative variance explained by the PCs

Figure 7.3 shows the biplot of the first two principal components (PC1 and PC2) from the PCA analysis. Each point represents a water quality parameter, and the lines indicate the direction and strength of the correlation with the principal components. According to this plot, temperature and DO exhibit similar patterns and are closely related in their behavior. BOD, COD, TSS, and turbidity are also strongly correlated, with turbidity and TSS being closely aligned, indicating similar behavior across the dataset. Other parameters show contrasting behavior compared to the first group.

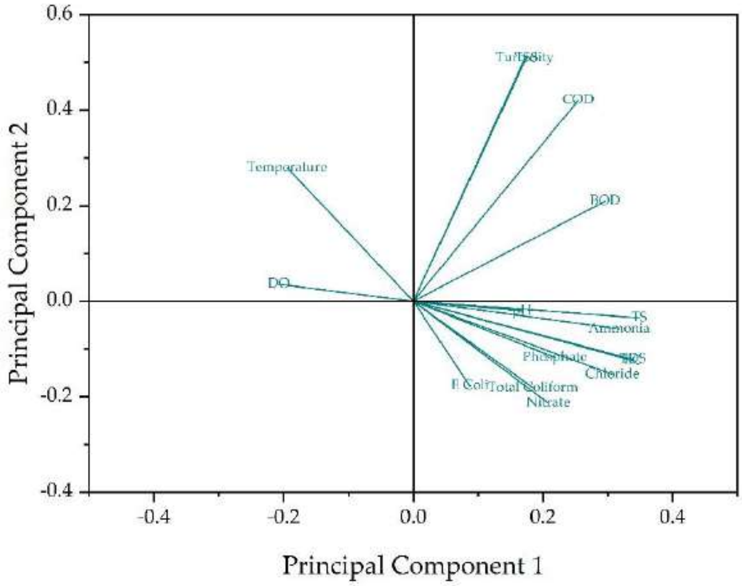


Figure 7.3: Biplot of the first two principal components

Parameters such as EC, TDS, TS, chloride, and ammonia-nitrogen exhibit high component loadings in the first principal component (PC). Turbidity, TSS, and COD show higher loadings in the second PC, while pH, DO, E. coli, and total coliform have higher loadings in the third PC. For temperature, BOD, nitrate, and phosphate, no PCs have a loading that exceeds 0.3. Consequently, these four parameters were excluded, and the remaining twelve parameters were retained for further analysis. The absolute loadings for each parameter in the PCs are shown in Figure 7.4.

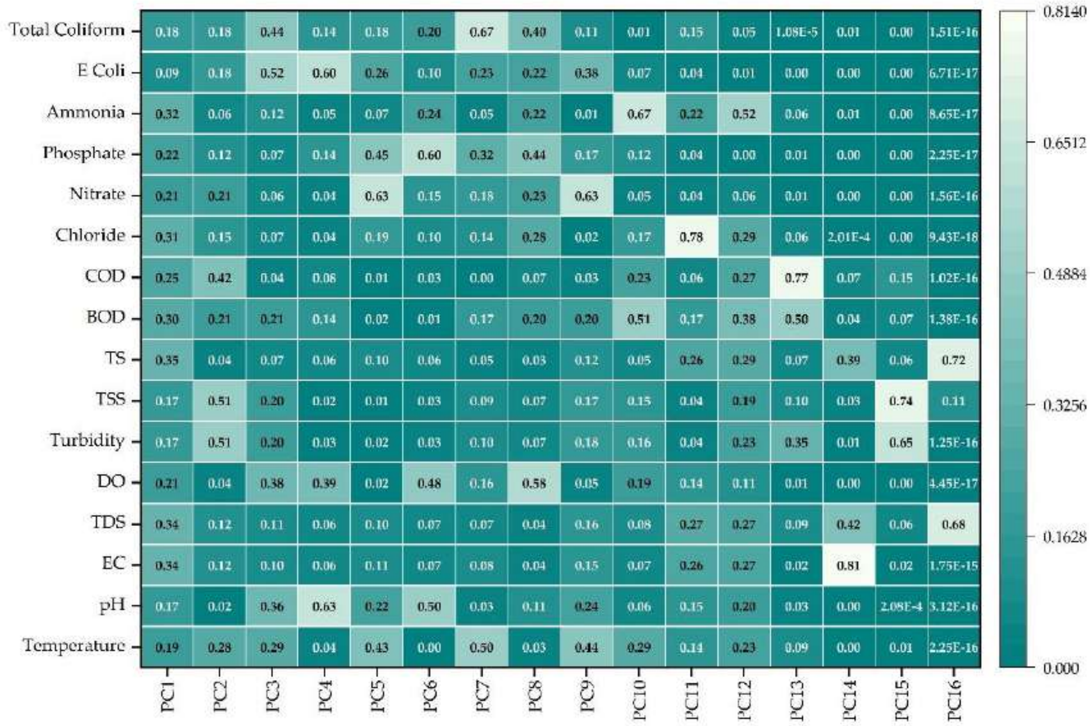


Figure 7.4: Absolute component loading

According to (Fartas et al., 2022), PCA combined with correlation analysis can be useful for reducing the number of parameters to include in the Water Quality Index (WQI) model. A similar approach was used in (Tripathi & Singal, 2019), where 28 water quality parameters were initially considered, reduced to 13 after PCA, and further reduced to 9 based on correlation analysis. In this study, correlation analysis was performed on the twelve water quality parameters selected from PCA to identify highly correlated parameters. The correlation matrix is shown in Figure 7.5. Two groups of parameters with a correlation coefficient greater than 0.9 were identified.

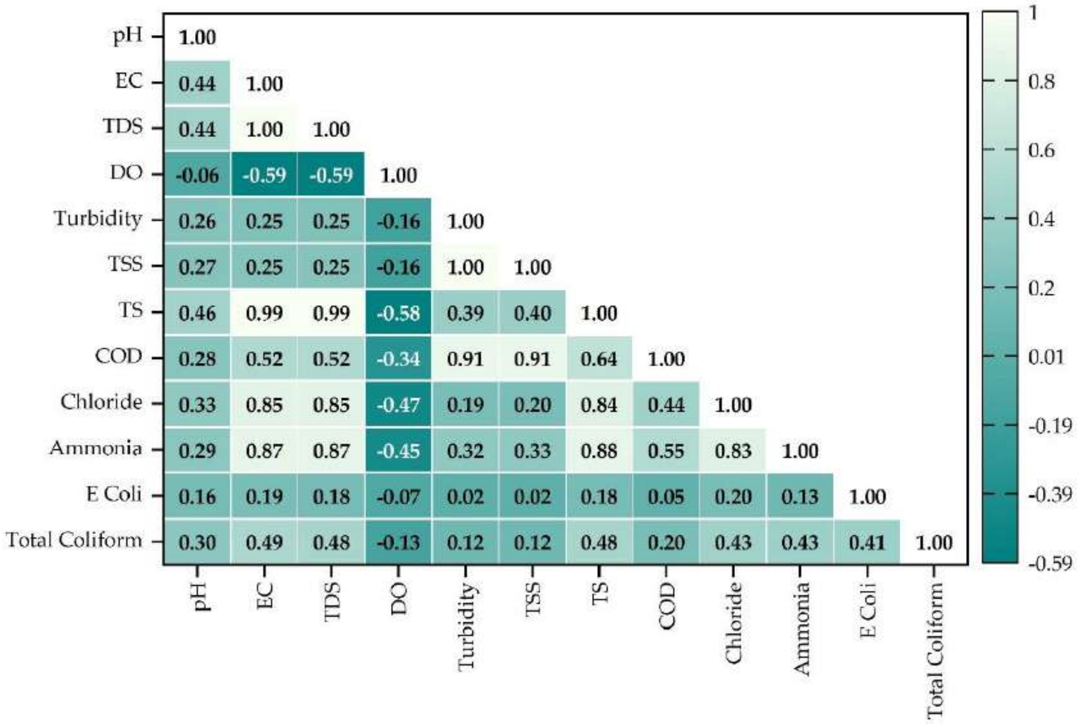


Figure 7.5: Correlation matrix of the selected parameters from PCA

The first group includes EC, TDS, and TS, while the second group includes COD, TSS, and turbidity. From these groups, TS and COD were retained. While EC measures the ability to conduct electricity, which correlates with the concentration of dissolved ions (Rusydi, 2018), TS provides a more direct measurement of solid content, including both dissolved and particulate matter. COD measures both organic and inorganic pollutants that affect oxygen levels in water (Gholizadeh et al., 2016). Therefore, it provides a more comprehensive and detailed assessment of water quality compared to TSS and turbidity. The other four parameters from those groups were discarded. Therefore, a total of 8 water quality parameters were selected for the WQI model.

7.1.2 Sub-indexing

The sub-indexing process is used to transform parameter concentrations into dimensionless values, referred to as the parameter sub-indices (Abbasi & Abbasi, 2012b). Sub-indices are determined using the values of WQPs along with quality rating curve functions. These rating curves relate the quality rating to the original parameter value. A quality rating for a specific parameter can be derived from its corresponding quality curve, known as the sub-index for that particular value (Nath Roy et al., 2024). Sub-index functions were created by using the water quality standards that help manage the river quality. These standards are based on different water uses and provide useful information for managing water (House, 1989), (S.-M. Liou et al., 2004). In this study, for the parameters that are also used in the original NSF-WQI, sub-indices were obtained from the rating curve function provided by the NSF-WQI model (Rai et al., 2012). The NSF model used suggested parameter ranges from water quality standards to calculate the sub-index values in a linear way (Lobato et al., 2015b), (Effendi et al., 2015b). For other parameters, not present in the NSF-WQI, simple linear interpolation functions were used. For this, the following equation was used (Sutadian et al., 2018b).

$$S_i = S_1 - \left[(S_1 - S_2)\left(\frac{X_1 - X_i}{X_1 - X_2}\right)\right]$$

where S_i is the sub-index value for water quality parameter i computed for the measured value X_i. S_1 and S_2 are the maximum and minimum sub-index values for the maximum and minimum guideline values (X_1 and X_2) for parameter i.

In the formulation, the thresholds for defining sub-index functions were established based on the permissible limits corresponding to different levels of intended uses or water quality classes. Five different water classes for sub-index values were established based on varying thresholds. However, in cases where specific thresholds for certain parameters were not available in Bangladesh's surface water quality standards, water quality standards from other countries, as well as those from international organizations such as the (FAO, 2023) and (WHO, 2024), were used. Additionally, similar approaches have been adopted in studies by (Stoner, 1978) , (S.-M. Liou et al., 2004), and (Thi Minh Hanh et al., 2011).

While developing the sub-indices by using this equation, a non-zero value for the lower end of each sub-indices were used instead of a value of 0. This was proposed by (House, 1989), as water will always has an economic value. Each parameter received a sub-index value of 100 for the best case and 5 for the worst case in this work.

Among the eight selected parameters, pH, DO, and TS are also included in the original NSF-WQI. Sub-indices for these parameters were derived using the rating curve function provided by the NSF-WQI model (Rai et al., 2012). The rating curves for these parameters are given in the Figure 7.6.

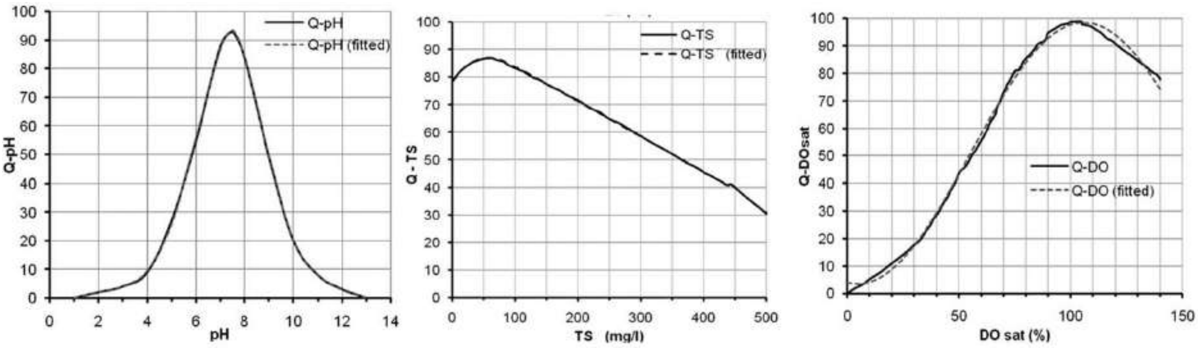


Figure 7.6: Rating curve for the parameters pH, total solids, and dissolved oxygen provided by NSF WQI

For the remaining five parameters, linear interpolation functions were applied. Table 3 presents the range of water quality parameters and their five key points used to define the rating curves, based on the standards proposed by (DoE, 2024), (DPHE, 2019), (ECR, 2023), (FAO, 2023), (WHO, 2024), and (Thi Minh Hanh et al., 2011).

To do so, surface water is classified into five distinct classes based on its suitability for various uses. These classes were defined based on designated uses of surface water, following a classification system from the literature. Class 1 represents surface water suitable for domestic water supply. Class 2 includes water that can be used for domestic supply with appropriate treatment or for the protection of aquatic life. Class 3 refers to water appropriate for irrigation purposes. Class 4 covers water that is usable for other activities requiring lower quality, such as navigation. Finally, Class 5 indicates wastewater that may be discharged into permitted water bodies for further treatment.

Table 7.1: Range of water quality parameters and defined score values for rating curves

Parameter	**Unit**	**Score Value**				
		Class 1	**Class 2**	**Class 3**	**Class 4**	**Class 5**
		100	**75**	**50**	**25**	**5**
COD	mg/L	10	15	30	50	80
Ammonia-Nitrogen	mg/L	0.1	0.2	0.5	1.0	10
Chloride	mg/L	250	400	600	-	-
E. coli	MPN/100 mL	20	50	100	200	-
Total Coliform	MPN/100 mL	2500	5000	7500	10000	-

The rating curves for these parameters are provided in the Figure 7.7.

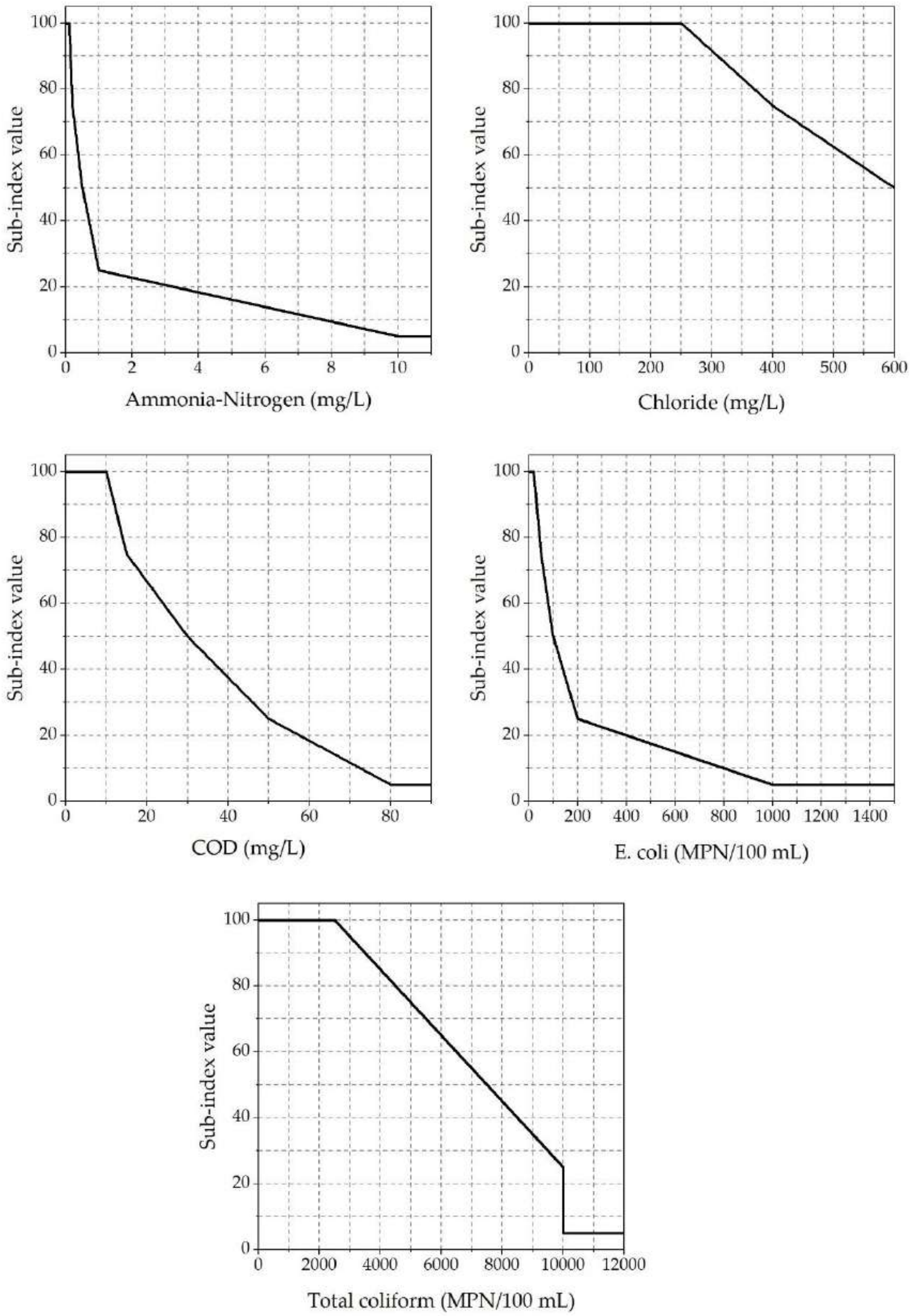


Figure 7.7: Rating curves generated by using linear interpolation function

7.1.3 Parameter Weight Calculation

The parameter weight value represents the relative importance of each selected water quality parameter (Sarkar & Abbasi, 2006a). In this study, unequal weighting techniques were employed, where the sum of all parameter weight values equaled 1. The weights for each parameter were calculated by using PCA result which method is proposed by (Naik et al., 2022) and (Sabinaya et al., 2024). PCA was performed again with the selected parameters and PCs having eigenvalue greater than 1 were selected. The eigenvalues of the selected PCs were then normalized and highest absolute parameter loading were taken for each parameter. Finally, the parameter weights were obtained by multiplication of normalized eigenvalue and the relative parameter loadings in the same PCs.

To calculate the weight of each parameter in the WQI model, PCA was performed again using the selected 8 parameters. From the outcome of the PCA, the first two principal components (PCs) were selected based on their eigenvalues being greater than 1. PC1 explains 49.35% variance and PC2 explains 15.82% variance. These two PCs explain 65.17% cumulative variance of overall dataset. Figure X shows the scree plot of the PCA results, where the eigenvalues of each PC are represented by the bars, and the line indicates the cumulative variance explained by the PCs.

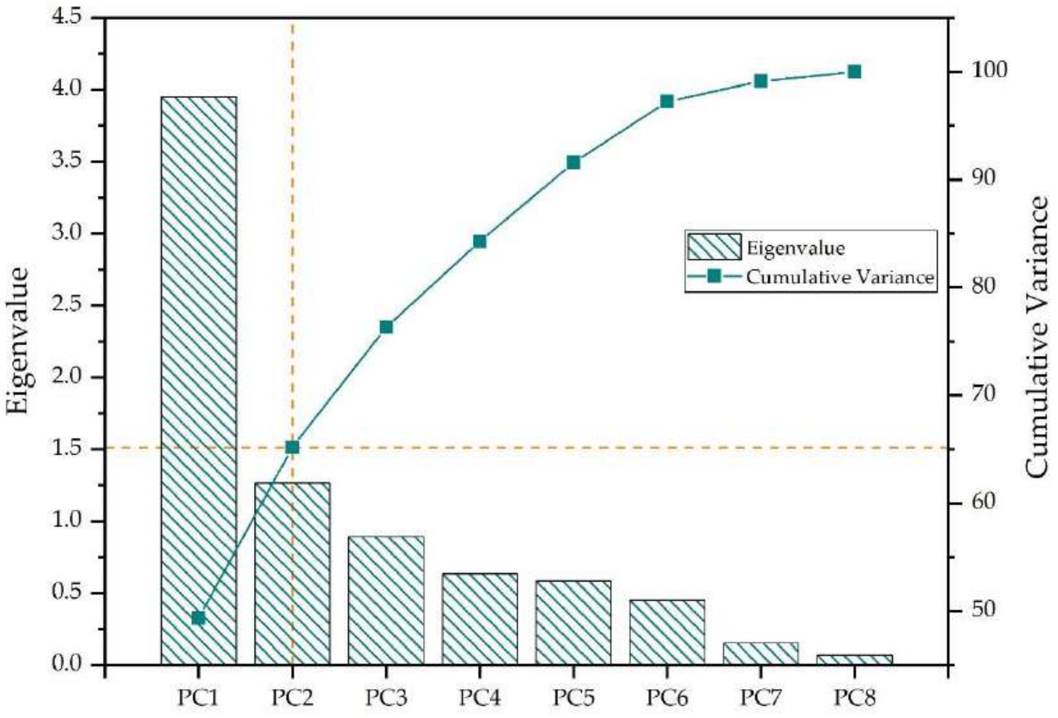


Figure 7.8: Scree plot showing eigenvalues and cumulative variance explained by the PCs

The biplot of the first two principal components (PC1 and PC2) from the PCA analysis is shown in Figure 7.9. According to this plot, pH, E coli, and total coliform exhibit similar patterns and are closely related in their behavior. Chloride, ammonia-nitrogen, total solids, and COD are also strongly correlated, with chloride, ammonia-nitrogen, and total solids being closely aligned, indicating similar behavior across the dataset. Dissolved oxygen shows contrasting behavior compared to the first group.

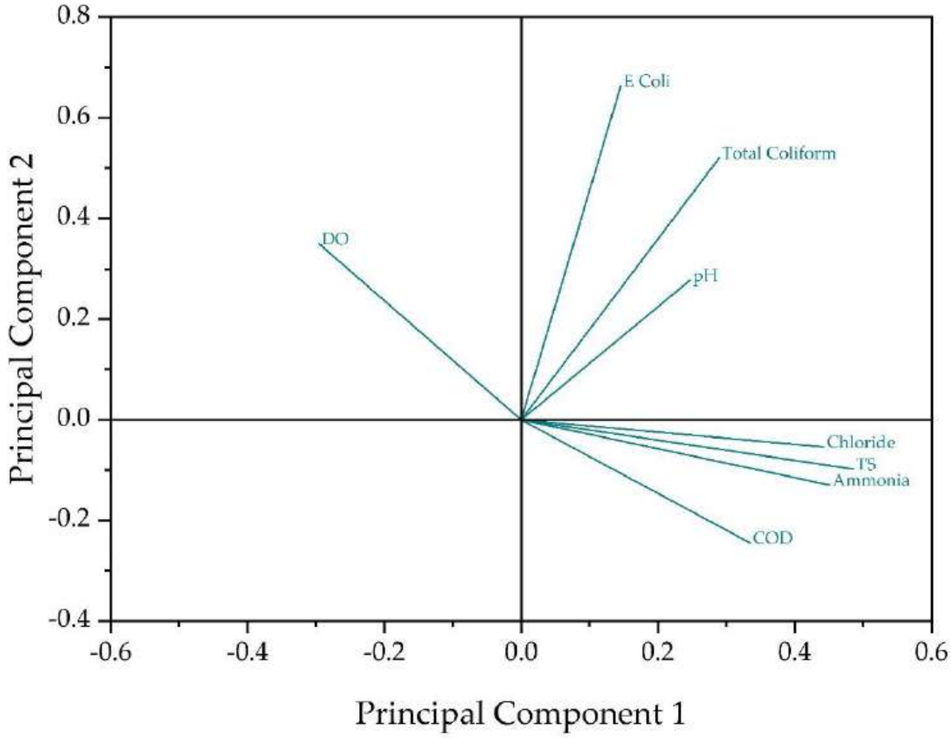


Figure 7.9: Biplot of the first two principal components

Considering the loading of parameters in the first two PCs, TS, COD, chloride, and ammonia-nitrogen have the highest absolute loadings in the first PC, while pH, DO, E. coli, and total coliform have the highest absolute loadings in the second PC. The absolute loading for each parameter is shown in the Figure 7.10.

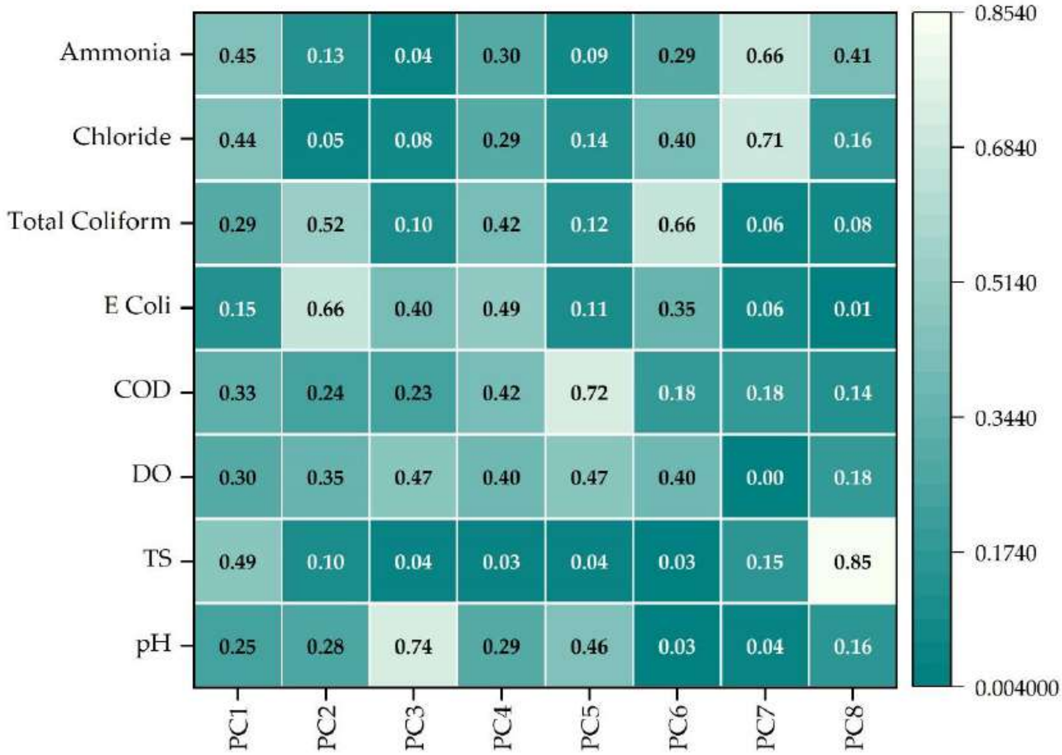


Figure 7.10: Absolute component loading

To calculate the weights, at first, the relative eigenvalue of the first two PCs were calculated. Then, the relative loading of each parameter in the same PC was multiplied by the relative eigen value of that PC. The final weights of the parameters are presented in Table 7.2.

Table 7.2: Assigned weights for the selected parameters

PC	Eigen Value	Relative eigen value (E_r)	Parameter	Highest absolute loading	Relative loading on same PC (L_r)	Weight ($E_r \times L_r$)
1	3.948	0.757	COD	0.335	0.195	0.148
			TS	0.486	0.284	0.215
			Chloride	0.443	0.258	0.195
			NH_3-N	0.450	0.263	0.199
			Total		1	
2	1.265	0.243	pH	0.277	0.153	0.037
			DO	0.350	0.193	0.047
			E. coli	0.664	0.367	0.089
			Total coliform	0.520	0.287	0.070
			Total		1	1

7.1.4 Aggregating Function and Classification

The aggregating function of the Water Quality Index (WQI) is a mathematical approach used to combine multiple water quality parameters into a single, composite index. It simplifies complex water quality data into an easy-to-understand score, typically ranging from 0 (poor quality) to 100 (excellent quality). To calculate the final WQI, a weighted average aggregating function was used. The mathematical expression of the aggregating function is denoted by the following equation:

$$WQI = \sum_{i=1}^{n} W_i Q_i$$

Where, W_i is the weight and Q_i is the sub-index value for the parameter i.

The weighted average aggregating function of the Water Quality Index (WQI) involves calculating a sub-index for each water quality parameter and multiplying it by its assigned weight. These weights reflect the relative importance of each parameter to overall water quality. The weighted sub-indices are then summed to produce the final WQI score.

The classification of WQI values is used to categorize water quality based on the final WQI score. Typically, WQI values are divided into several ranges to indicate the quality of water, such as excellent, good, fair, poor, and very poor. These classifications help in understanding the suitability of water for various uses, such as drinking, irrigation, or recreation. The exact thresholds for each category may vary depending on the specific WQI model, but they generally provide a clear and standardized way to assess and communicate water quality. In this newly developed WQI model, the water quality is classified based on the WQI values according to the classification proposed by NSF-WQI. The NSF-WQI classification is shown in the Table 7.3.

Table 7.3: Classification of WQI value

WQI range	Category	Color code
90-100	Excellent	
70-89	Good	
50-69	Medium	
25-49	Bad	
0-24	Very Bad	

7.2 Application of the Developed WQI Model

The application of the developed WQI model is demonstrated through the evaluation of water quality based on data collected from four rivers across four different seasons. Each river has nine sampling stations, and the WQI is evaluated for each sample station in every season. The WQI model integrates multiple water quality parameters into a single index value, providing a comprehensive assessment of water quality. By applying this model to the collected data, the analysis offers a clear and objective means to evaluate the suitability of water for various uses in different locations and seasonal conditions, highlighting the effectiveness of the WQI in simplifying complex water quality data into an actionable format.

7.2.1 Visual Representation of Seasonal WQI Variations

The evaluation of the Water Quality Index (WQI) for the Buriganga River revealed significant seasonal variations in water quality across different sampling stations. The minimum WQI value recorded was 30.50 at the Chandir Ghat station during the dry season, falling into the "Bad" category. Similarly, all sampling stations in the dry and pre-monsoon seasons exhibited WQI values in the "Bad" category, although a slight improvement was observed in the pre-monsoon season compared to the dry season. Water quality showed notable improvement during the monsoon and post-monsoon seasons. Figure 7.11 shows the WQI values at each sampling stations.

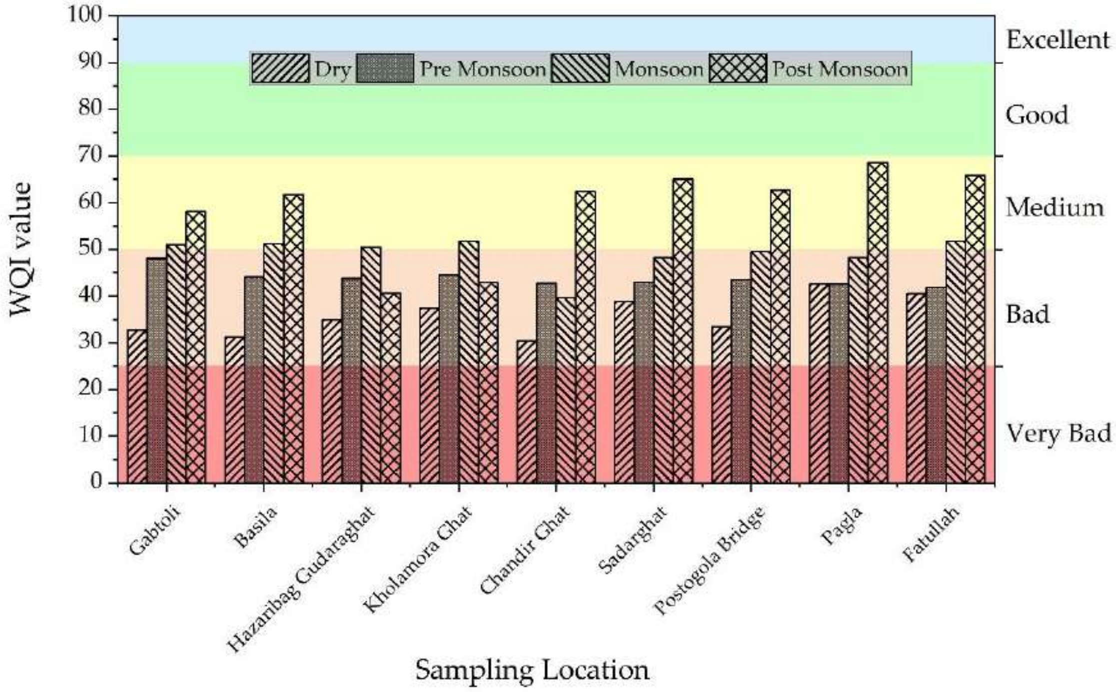


Figure 7.11: WQI values in each sampling stations of Buriganga River

In the monsoon, the upper stream stations exhibited "Medium" category water quality, while in the post-monsoon season, all stations except Hazaribag Gudaraghat and Kholamora Ghat showed water quality in the "Medium" category. Hazaribag Gudaraghat and Kholamora Ghat remained in the "Bad" category. The overall trend indicates an improvement in water quality over time, with the exception of Hazaribag Gudaraghat, Kholamora Ghat, and Chandir Ghat, where fluctuations were observed. These findings highlight both seasonal and locational variations in water quality along the Buriganga River.

The assessment of the WQI for the Balu River highlights significant seasonal changes in water quality across various sampling stations. The minimum WQI value recorded was 30.52 at the Talia Termukh Bridge station during the dry season, categorizing the water quality as "Bad." Similarly, during both the dry and pre-monsoon seasons, all WQI values fell within the "Bad" category, though a slight increase in WQI was observed in the pre-monsoon season compared to the dry season. In the monsoon and post-monsoon seasons, water quality improved significantly. During the monsoon, the upper stream stations, Pubail Bridge and Ulukhola Bazar, exhibited "Good" water quality, while all other stations fell into the "Medium" category. In the post-monsoon season, although the WQI decreased at Pubail Bridge, Ulukhola Bazar and Talia Termukh Bridge showed "Medium" water quality. Water quality generally improved over time, with the exception of Pubail Bridge and Ulukhola Bazar. At Pubail Bridge, the WQI decreased from the monsoon to the post-monsoon season, while at Ulukhola Bazar, the WQI remained stable in the post-monsoon. These findings indicate both seasonal and locational variations in water quality across different stations along the Balu River, highlighting a general trend of water quality improvement over time, particularly in the monsoon and post-monsoon seasons. Figure 7.12 shows the WQI values at each sampling stations.

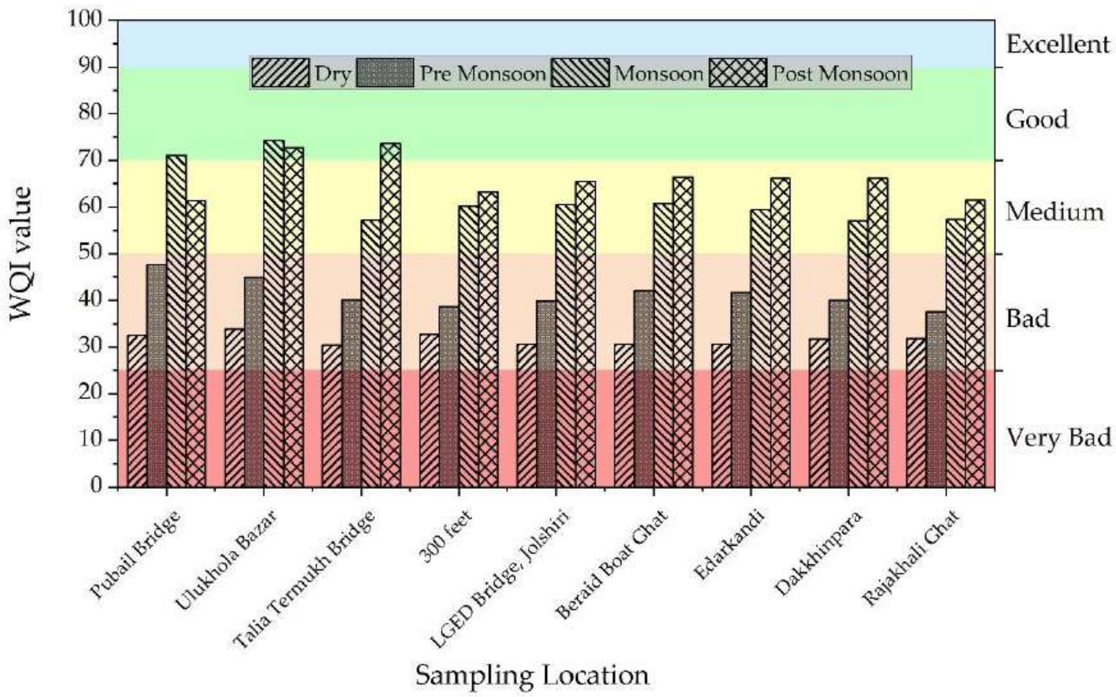


Figure 7.12: WQI values in each sampling stations of Balu River

The minimum WQI value recorded was 29.78 at Fulpukuria Thread during the dry season, reflecting "Bad" water quality, and during this season, all stations fell into the "Bad" category. However, starting in the pre-monsoon season, water quality began to improve. In the pre-monsoon, monsoon, and post-monsoon seasons, the majority of stations recorded "Medium" water quality, with a few exceptions. The maximum WQI value of 73.12 was observed at Rustompur Ghat during the monsoon season, indicating "Good" water quality at this location. In the post-monsoon season, downstream stations Ashulia Landing Station, Rustompur Ghat, and Birulia showed "Good" water quality. While water quality was much worse in the dry season compared to other seasons, it generally improved over time. However, this improvement was not consistent at all stations, with some stations showing fluctuations in water quality across the seasons. Figure 7.13 shows the WQI values at each sampling stations.

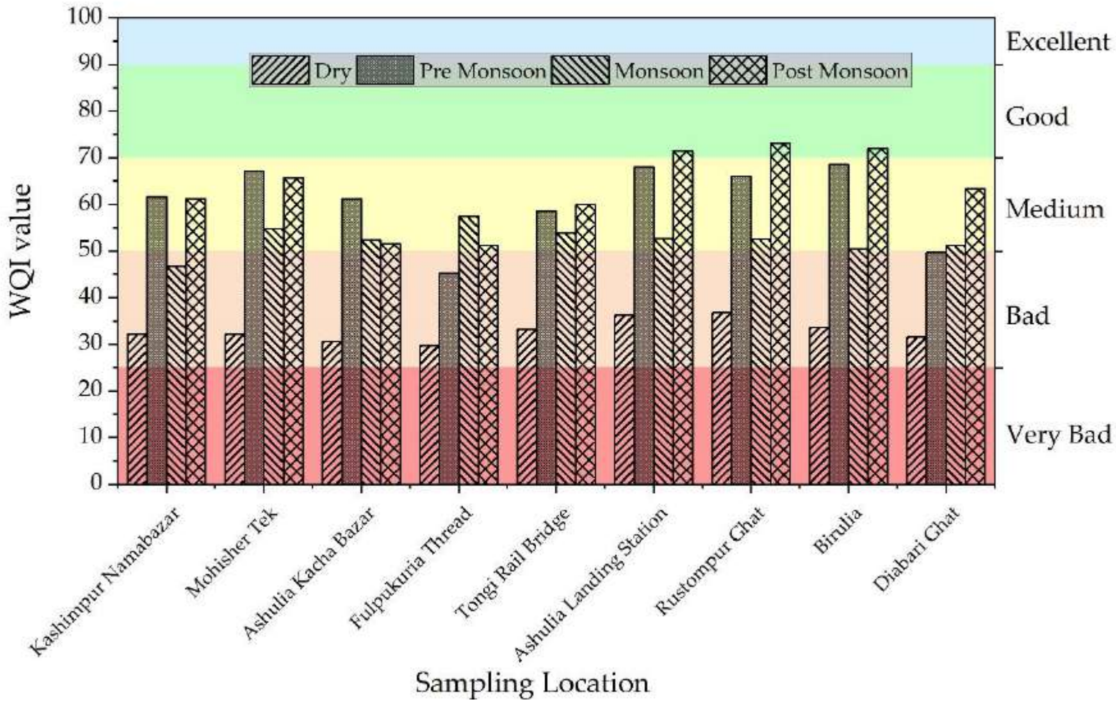


Figure 7.13: WQI values in each sampling stations of Turag River

The minimum WQI value recorded for the Sitalakhya River was 36.49 at Horipur Power Plant during the pre-monsoon season, while the maximum WQI value was 77.96 at Beldi Bazar during the post-monsoon season. In the dry season, the upper stream stations exhibited "Medium" water quality, while the water quality in the downstream areas significantly degraded, with stations falling into the "Bad" category. During the pre-monsoon season, water quality worsened compared to the dry season, with almost all stations falling into the "Bad" category, except for Kaliganj and Beldi Bazar, which maintained better water quality. In the monsoon season, water quality improved across all stations, with all of them falling into the "Medium" category. The most notable improvement occurred in the post-monsoon season, where the water quality at the upstream stations improved further. Five upstream stations, excluding Hatabo Bazar, showed "Good" water quality, while the remaining stations in the downstream areas fell into the "Medium" category. Throughout all seasons, the upper stream locations of the Sitalakhya River consistently exhibited better water quality than the downstream areas. Figure 7.14 shows the WQI values at each sampling stations.

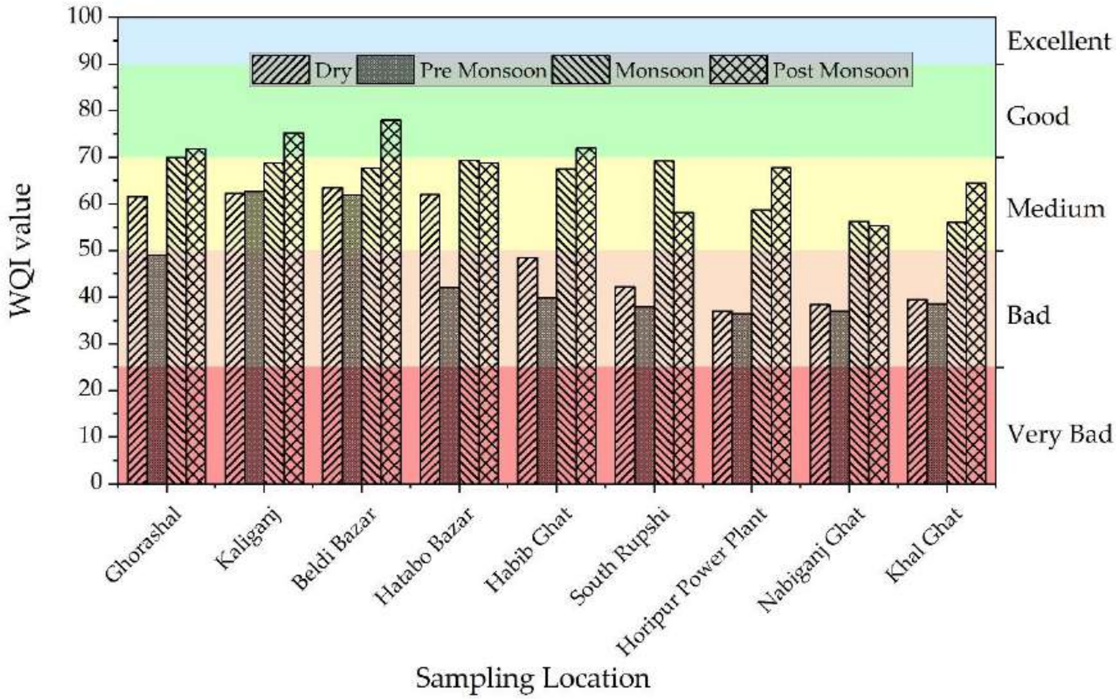


Figure 7.14: WQI values in each sampling stations of Sitalakhya River

The overall distribution of water quality index values across different categories for the collected dataset are given in the Table 5. Out of a total of 144 data points, the majority fall within the "Medium" and "Bad" categories, with 67 entries (46.53%) in the "Medium" range and 66 entries (45.83%) in the "Bad" range. A small portion, 11 entries (7.64%), fall within the "Good" category. There are no data points in the "Excellent" or "Very bad" categories. This suggests that the water quality is predominantly in the medium to bad range, with very few instances of good quality.

Table 7.4: Overall WQI distribution for the collected water quality data

WQI Value	Remarks	Count	Percentage (%)
90-100	Excellent	0	0
70-89	Good	11	7.64
50-69	Medium	67	46.53
25-49	Bad	66	45.83
0-24	Very bad	0	0
Total		144	100

7.2.2 GIS-Based Visualization of River Water Quality

The spatial distribution of Water Quality Index (WQI) values was visualized using ArcGIS through interpolation techniques to provide a clearer understanding of spatial variations across sampling locations (Figure 7.15). This GIS-based mapping approach allows the identification of pollution hotspots and regional differences in water quality, supporting effective planning and localized interventions. The WQI values were categorized and color-coded according to water quality classes, making it easier to interpret the conditions of different river stretches.

During the dry season, the map shows that all sampling stations along the Buriganga, Balu, and Turag rivers, as well as the downstream section of the Sitalakhya River, fall within the "bad" water quality category. The only exception is the upper stream of the Sitalakhya River, which maintains a "medium" water quality status. These poor conditions are likely due to reduced river flow, which limits the dilution of untreated domestic and industrial discharges. Persistent pollution from textile, tannery, and other industrial sources, along with municipal waste, contributes to the degraded conditions.

In the pre-monsoon season, while water quality in the Buriganga River remains within the "bad" category, a notable improvement is observed in the Turag River, where most stations shift to the "medium" category—except for Fulpukuria Thread, which remains "bad." This improvement may result from slightly increased flow and reduced pollutant discharge. However, Balu River and the downstream Sitalakhya continue to exhibit poor. A slight degradation in water quality is observed in the upper Sitalakhya, possibly due to increased runoff from surrounding urban areas and nearby discharges.

During the monsoon season, most stations shift to the "medium" category, indicating a general improvement likely due to the flushing effect of increased rainfall and river discharge. However, several downstream stations on the Buriganga River—including Chandir Ghat, Sadar Ghat, Postogola, and Pagla—remain in the "bad" category. These areas are characterized by high population density, heavy river traffic, and continuous discharge of untreated waste. The upstream station of the Turag River also falls within the "bad" category, possibly due to localized point-source pollution or backflow from connected drainage systems.

In the post-monsoon season, further improvements are evident. The upper reaches of the Sitalakhya River reach the "good" category, suggesting effective natural recovery and dilution after the monsoon. Some stations on the Balu and Turag rivers also improve to the "good" category. Notable examples include Ulukhola Bazar and Talia on the Balu, and Birulia, Rustompur Ghat, and Ashulia Landing on the Turag. These improvements may be attributed to the sustained high flow during the monsoon and reduced pollutant inputs afterward.

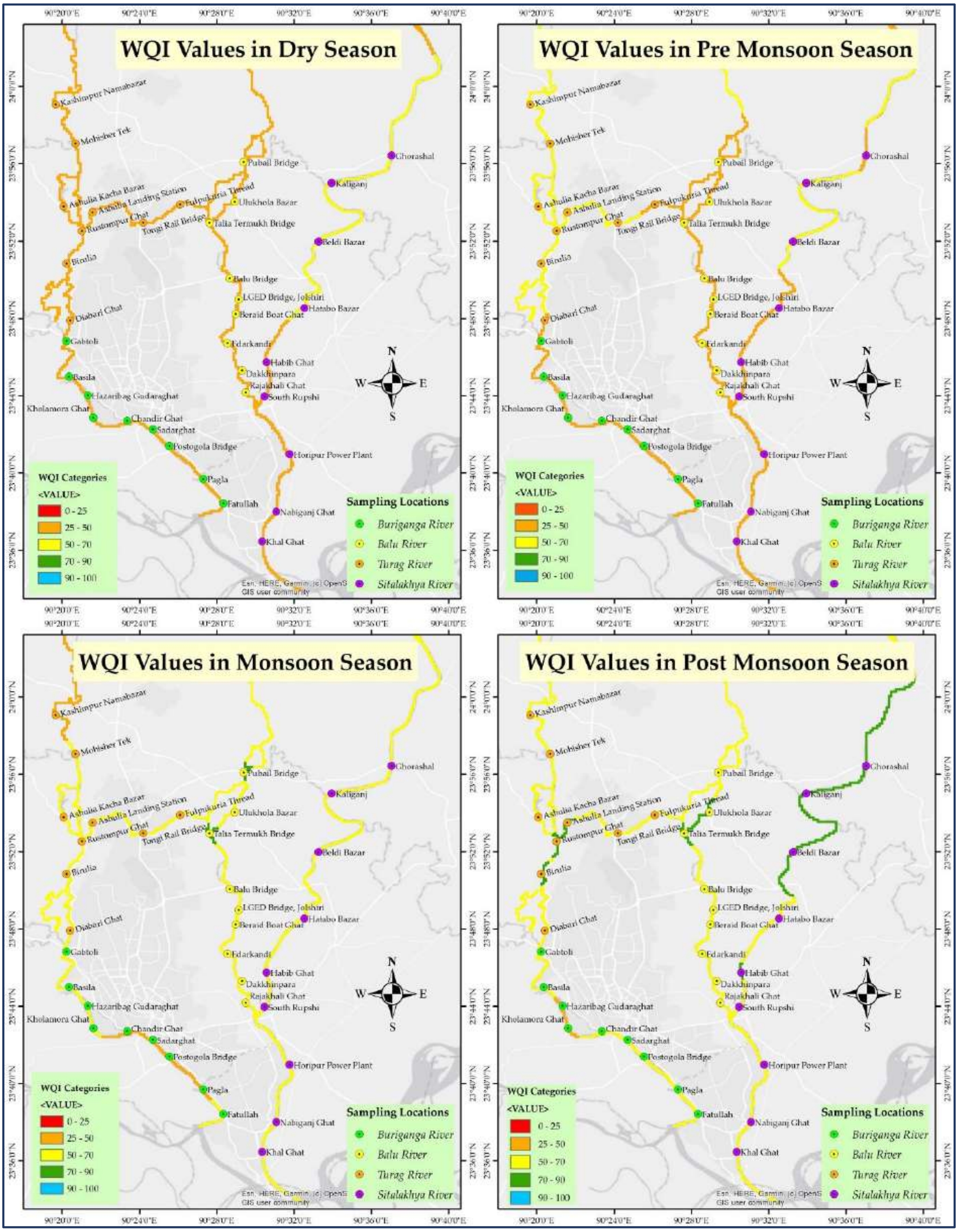


Figure 7.15: WQI values along the rivers

7.2.3 Statistical Summary of WQI Values

A box-and-whisker plot is a graphical representation of data that displays its minimum, first quartile (Q1), median, third quartile (Q3), and maximum. From it, the central tendency, spread, and any potential outliers in the dataset are identified, and the distribution's symmetry or skewness is observed. For the evaluated WQI values, a box-and-whisker plot can be used to visually summarize these characteristics, providing insight into the data's overall distribution and identifying any irregularities or patterns.

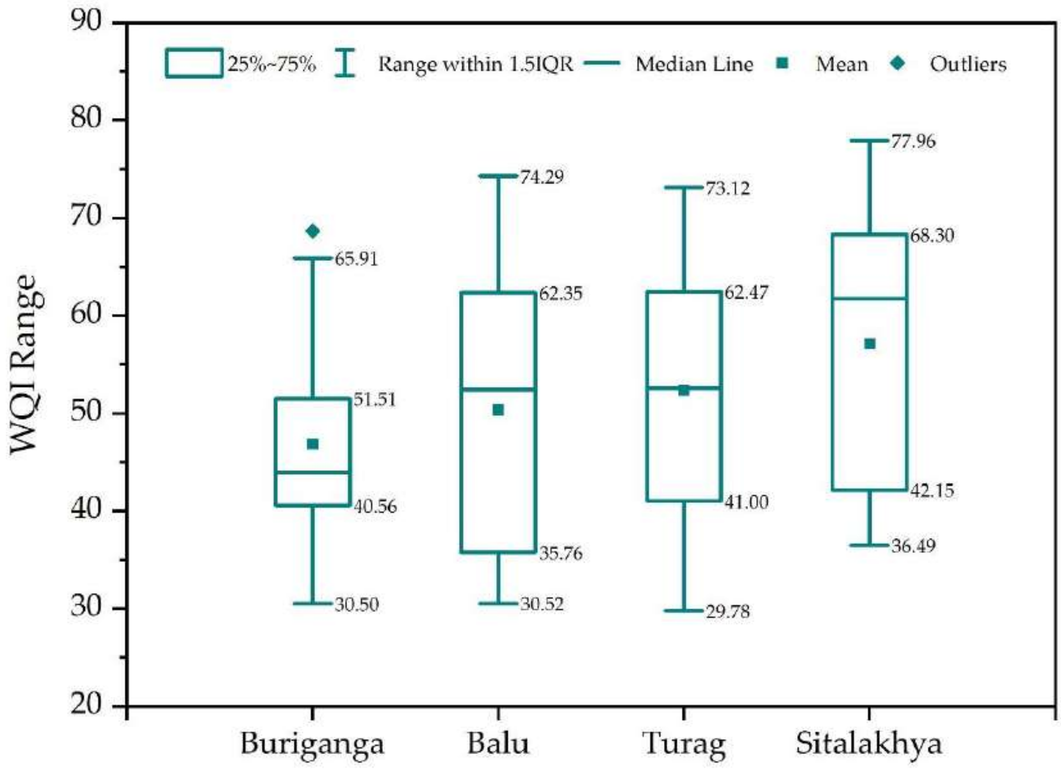


Figure 7.16: Box-and-whisker plot of the WQI data for the year 2024

The box-and-whisker plot, shown in the Figure 7.16, of the WQI data for Rivers Buriganga, Balu, Turag, and Sitalakhya highlights distinct differences in water quality. Buriganga River, with a range from 30.50 to 68.65 and a median of 43.94, shows the lowest overall water quality. The mean WQI value is 46.85, which is also the lowest among the rivers, and the presence of an outlier at 68.65 suggests occasional spikes in water quality, though it generally remains poor. Balu River has a WQI range from 30.52 to 74.29, with a median of 52.42 and a mean of 50.35, indicating moderate water quality. Turag River, with a range of 29.78 to 73.12, has a median of 52.56 and a mean of 52.33, suggesting a slightly better and more consistent water quality than the Balu River. Sitalakhya River, however, demonstrates the best water quality, with a range from 36.49 to 77.96, a median of 61.74, and a mean of 57.10, indicating consistently higher water quality throughout the year. Overall, Turag and Sitalakhya Rivers exhibit superior water quality compared to Buriganga and Balu Rivers, with Sitalakhya showing the most stable and highest values.

7.3 Sensitivity and Uncertainty Analysis

A global variance-based Sobol sensitivity analysis was conducted to assess the relative importance of various water quality parameters in influencing the variability of the Water Quality Index (WQI). Sobol sensitivity analysis is a powerful technique used to decompose the variance of a model's output into contributions from individual input parameters and their interactions, providing insights into which factors most significantly impact the model's behavior. This approach is particularly useful when dealing with complex models where the relationships between parameters and outcomes may be nonlinear or involve intricate interactions. The analysis involves the calculation of Sobol indices, which quantify the sensitivity of the output (WQI in this case) to each parameter by partitioning the total variance into components corresponding to individual parameters and their interactions.

The Sobol indices are categorized into first-order indices (Si) and total-order indices (STi). The first-order indices (Si) capture the direct, independent contribution of each parameter to the variance of the output. In contrast, the total-order indices (STi) account for both the direct contribution of a parameter and its interactions with other parameters. This distinction is crucial as it allows for the identification not only of the parameters with the most substantial direct influence on the output but also of those whose effect is amplified or modified through interactions with other variables.

Figure 7.17 presents the calculated Sobol indices for the selected water quality parameters. The first-order Sobol indices (Si) reveal that Total Solids (TS) (Si = 0.2506), Chemical Oxygen Demand (COD) (Si = 0.2274), and chloride (Si = 0.2013) are the most influential parameters, implying that these factors have the greatest direct effect on the variability of the WQI. The relatively high first-order indices for TS, COD, and chloride indicate that variations in these parameters alone can substantially explain changes in the WQI, highlighting their critical role in determining water quality. These findings are consistent with previous studies on urban water bodies, where TS and COD have been frequently identified as dominant water quality indicators due to their direct impact on the overall health of aquatic environments.

The total-order Sobol indices (STi) provide additional insight into the interactions between parameters. For example, dissolved oxygen (DO) and ammonia-nitrogen exhibit interaction effects, as evidenced by their total-order indices being higher than their first-order indices: 0.0516 vs. 0.0488 for DO and 0.1674 vs. 0.1669 for ammonia-nitrogen. This suggests that the variability in DO and ammonia-nitrogen concentrations is influenced not only by their individual values but also by their interactions with other parameters, indicating a degree of interdependence between these factors. The presence of such interaction effects underscores the complexity of the system, where the combined influence of parameters can alter the overall WQI in ways that are not captured by considering each parameter in isolation.

On the other hand, parameters such as pH (Si = 0.0224), Escherichia coli (E. coli) (Si = 0.0458), and total coliform (Si = 0.0365) exhibit minimal contributions to WQI variability, as indicated by their low first-order indices. These results suggest that these parameters have a limited direct impact on the WQI in the studied system, which may be due to their relatively stable concentrations or less significant role in shaping overall water quality in the specific context of the analysis.

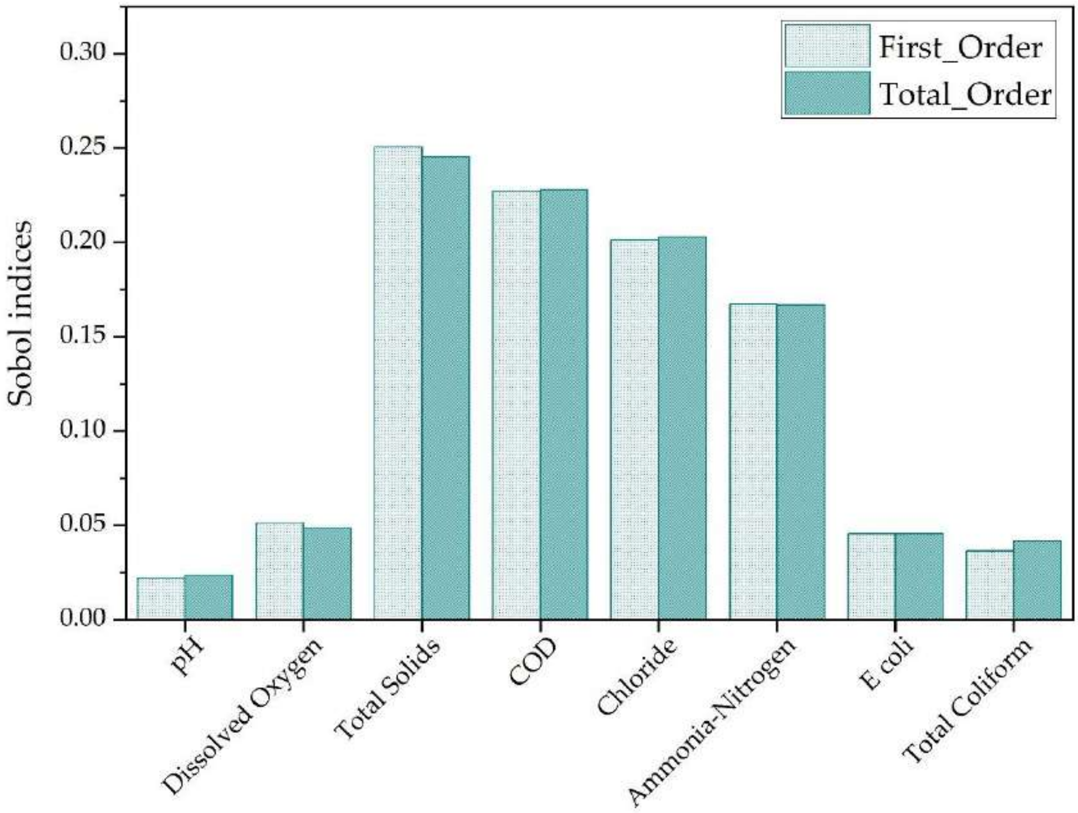


Figure 7.17: Sobol sensitivity indices for each parameter

In conclusion, the Sobol sensitivity analysis reveals that TS, COD, and chloride are the primary drivers of WQI variability, while DO and ammonia-nitrogen exhibit notable interaction effects. The results further suggest that certain parameters, such as pH, E. coli, and total coliform, have a minimal influence on the overall water quality as measured by the WQI. These findings provide valuable insights for water quality management, emphasizing the need to focus on the most influential parameters and consider their interactions when evaluating or modeling water quality in urban water bodies.

To quantify the uncertainty in WQI estimates, stratified bootstrap resampling was performed with 5000 iterations, generating confidence intervals (CIs) for each river. Bootstrapping is a non-parametric statistical technique that involves resampling the observed data with replacement to generate multiple simulated datasets. This approach provides robust estimates of uncertainty without assuming a specific underlying distribution for the data, which is particularly useful in environmental studies where data may not conform to a normal distribution. Stratification was applied to account for seasonal variations, helping to reduce seasonal bias in the resampling process and ensuring a more representative estimation of uncertainty. By assessing the variability across the resampled datasets, bootstrapping allows for a comprehensive evaluation of uncertainty in the WQI estimates. The results of the bootstrap analysis, including the mean WQI values and their associated 95% confidence intervals (CIs), are shown in Figure 7.18.

The Sitalakhya River exhibited the highest mean WQI among the four rivers, with a value of 48.49. The 95% confidence interval ranged from 46.29 to 50.64, and the standard deviation was 1.13. These results suggest relatively high water quality; however, the broader confidence interval and higher variability indicate a greater degree of uncertainty in the WQI estimate. This may reflect fluctuating environmental conditions or varying pollution loads affecting the river throughout the year.

The Turag River had a mean WQI of 44.40, with a 95% confidence interval from 42.90 to 45.95 and a standard deviation of 0.93. While the uncertainty level is moderate, the WQI distribution exhibited a bimodal pattern, suggesting the influence of distinct seasonal that result in two dominant water quality regimes. This pattern highlights the importance of accounting for temporal variability when assessing water quality in this river.

For the Buriganga River, the mean WQI was 42.33, with a 95% confidence interval ranging from 40.38 to 44.05 and a standard deviation of 1.01. The relatively wider confidence interval indicates a moderate level of uncertainty in the water quality estimate. This suggests that the Buriganga River experiences variability in water quality, potentially due to diverse pollution sources or seasonal fluctuations.

The Balu River reported the lowest mean WQI at 40.82, with a 95% confidence interval between 39.82 and 42.32 and the lowest standard deviation of 0.61. The narrow confidence interval and low variability indicate relatively stable water quality with less uncertainty in the WQI estimate compared to the other rivers.

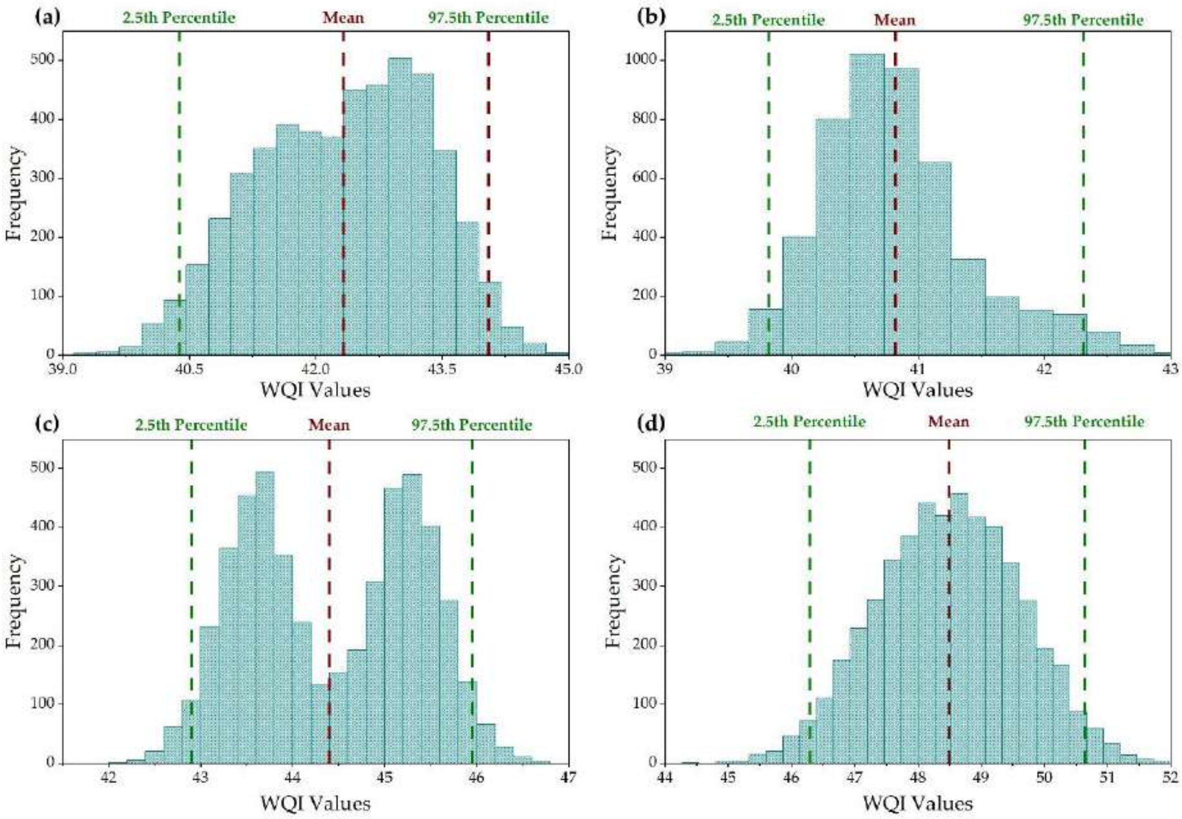


Figure 7.18: Bootstrap distribution of WQI for (a) Buriganga (b) Balu (c) Turag (d) Sitalakhya river

In comparison, the Sitalakhya River had the highest WQI but also showed the highest uncertainty, while the Balu River had the lowest WQI and the least uncertainty, indicating more consistent water conditions. The Buriganga and Turag rivers fell between these two extremes, both showing moderate uncertainty, though the Turag's bimodal distribution suggests more complex temporal dynamics. The use of stratified bootstrap resampling helped to minimize seasonal bias across all cases, leading to more balanced and representative uncertainty estimates in the WQI assessments.

The combined insights from sensitivity and uncertainty analyses provide practical guidance for improving monitoring and management strategies. Parameters with the highest influence on WQI variability, such as TS, COD, and chloride should be prioritized for targeted monitoring to enhance the model's responsiveness to pollution changes. In contrast, the high uncertainty observed in the Sitalakhya River highlights limitations in model reliability for certain locations, emphasizing the need for more frequent sampling and site-specific assessments. The bimodal pattern of WQI in the Turag River indicates the need for improved temporal resolution of sampling during transitional seasons, along with focused assessment of pollution sources. Addressing such uncertainty through regular monitoring and a focus on influential parameters is essential to support informed and reliable management decisions.

Chapter 8: Insights into Temporal Trends and Methodological Comparison of the Developed WQI

8.1 Comparison of the Developed WQI using 2019 and 2024 Data

Two distinct Water Quality Index (WQI) models were developed in this study to assess the surface water quality of selected rivers in two different time periods. The first model was constructed using water quality data from 2019, collected from the Department of Environment (DoE), while the second model was based on primary data collected in 2024. Both models were developed following a similar methodological framework but with different input parameters and data sources.

The two WQI models were constructed using different sets of water quality parameters. The 2019 WQI model was developed using seven parameters: pH, total alkalinity (TA), total solids (TS), dissolved oxygen (DO), biochemical oxygen demand (BOD), turbidity, and temperature. These parameters primarily reflect general physico-chemical characteristics and organic pollution indicators. In contrast, the 2024 WQI model included pH, DO, chemical oxygen demand (COD), TS, chloride, ammonia-nitrogen, E. coli, and total coliform, representing a more comprehensive focus on chemical contaminants and microbial pollution. The variation in parameter selection naturally influenced the structure and sensitivity of each model, with the 2024 WQI more directly capturing potential health-related risks.

Both the 2019 and 2024 datasets include sampling data from the Buriganga, Balu, Turag, and Sitalakhya rivers. To maintain consistency in spatial comparison, only the common sampling stations between the two periods were considered for analysis. The Buriganga River had the highest overlap, with data from seven identical stations available in both years. The Balu River had four stations, while the Turag and Sitalakhya rivers each had three. This station-level alignment allowed for a reliable spatial comparison of WQI values over time, reducing potential biases introduced by differing sampling locations and ensuring that observed changes more accurately reflect temporal variations in water quality rather than spatial discrepancies.

Figure 8.1 presents the WQI values for Buriganga River. In January and February of 2019, the WQI values at each sampling station were similar, indicating relatively stable water quality within that period at individual sites. However, in the dry season of 2024, WQI values decreased noticeably across most stations, reflecting a deterioration in water quality over time. This decline suggests increasing environmental pressures or pollution impacts during the dry season between the two timeframes. During the pre-monsoon season of 2024, the WQI values were generally lower than those recorded in March and May of 2019. Notably, the WQI in April 2019 was lower compared to both March and May 2019, indicating a mid-season decline in water quality that year. The WQI values for the pre-monsoon period in 2024 were closely comparable to the April 2019 values, suggesting a similar level of water quality during these times. During the monsoon season, water quality declined significantly in 2024. In June through September of 2019, the WQI values remained relatively stable across all sampling stations, showing little fluctuation and a consistent range. However, in the 2024 monsoon season, the WQI values decreased markedly, indicating a substantial deterioration in water quality compared to 2019. In October 2019, the WQI values were higher compared to those in

November 2019 and were similar to the values observed during the 2019 monsoon season. In contrast, during the post-monsoon period of 2024, the WQI values were lower than those recorded in November 2019, indicating a further decline in water quality in the later year.

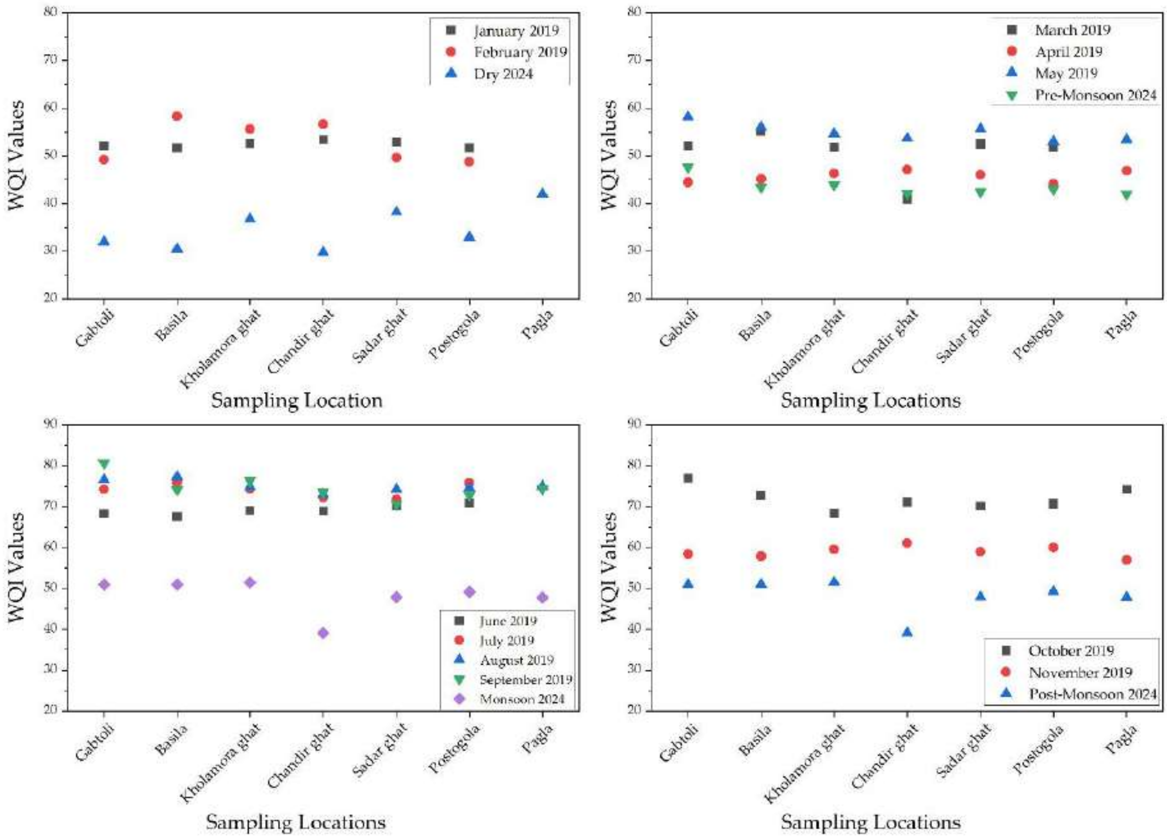


Figure 8.1: Comparison of WQI values across common sampling stations of Buriganga River

Figure 8.2 presents the WQI values for the Balu River. During the dry season of 2024, WQI values were generally lower than those recorded in January and February of 2019. While moderate variation in WQI was observed across most sampling stations, the Beraid station showed consistent WQI values in both January 2019 and the 2024 dry season. In the 2019 pre-monsoon season, May exhibited the highest WQI values, whereas April had the lowest within the same period. The 2024 pre-monsoon WQI values closely resembled those of April 2019. During the monsoon season, water quality in 2024 showed a slight deterioration compared to 2019. However, in the post-monsoon season, WQI values for both years were nearly identical.

Figure 8.3 presents the WQI values for the Turag River. For the Turag River, during the dry season in January and February of 2019, WQI values ranged between 50 and 60, with February showing slightly higher values. However, in 2024, the WQI values decreased significantly, falling to between 30 and 35. During the pre-monsoon season, the Diabari station recorded similar WQI values in April 2019 and pre-monsoon 2024, while the Ashulia Kachabazar station showed higher WQI values in 2024 compared to 2019. In the monsoon season, water quality in 2024 deteriorated significantly compared to 2019. During the post-monsoon period, the 2024 WQI values were slightly lower than those observed in 2019.

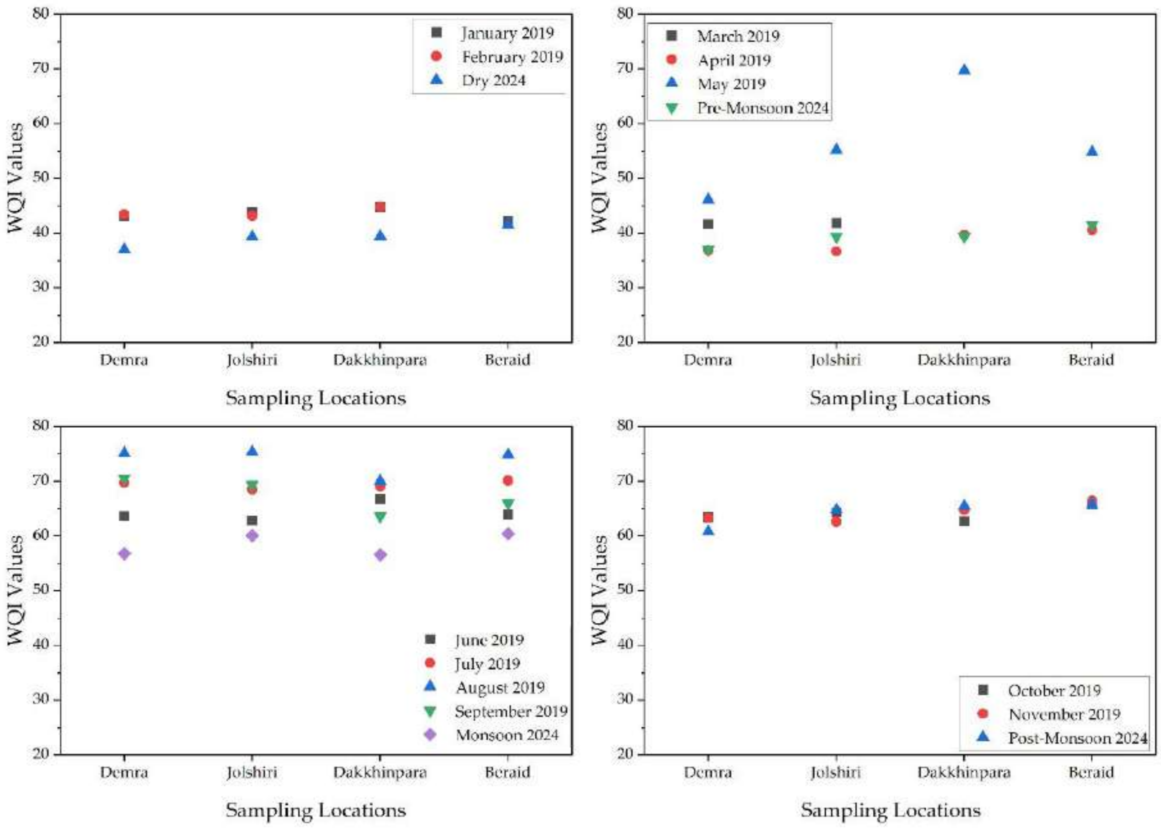


Figure 8.2: Comparison of WQI values across common sampling stations of Balu River

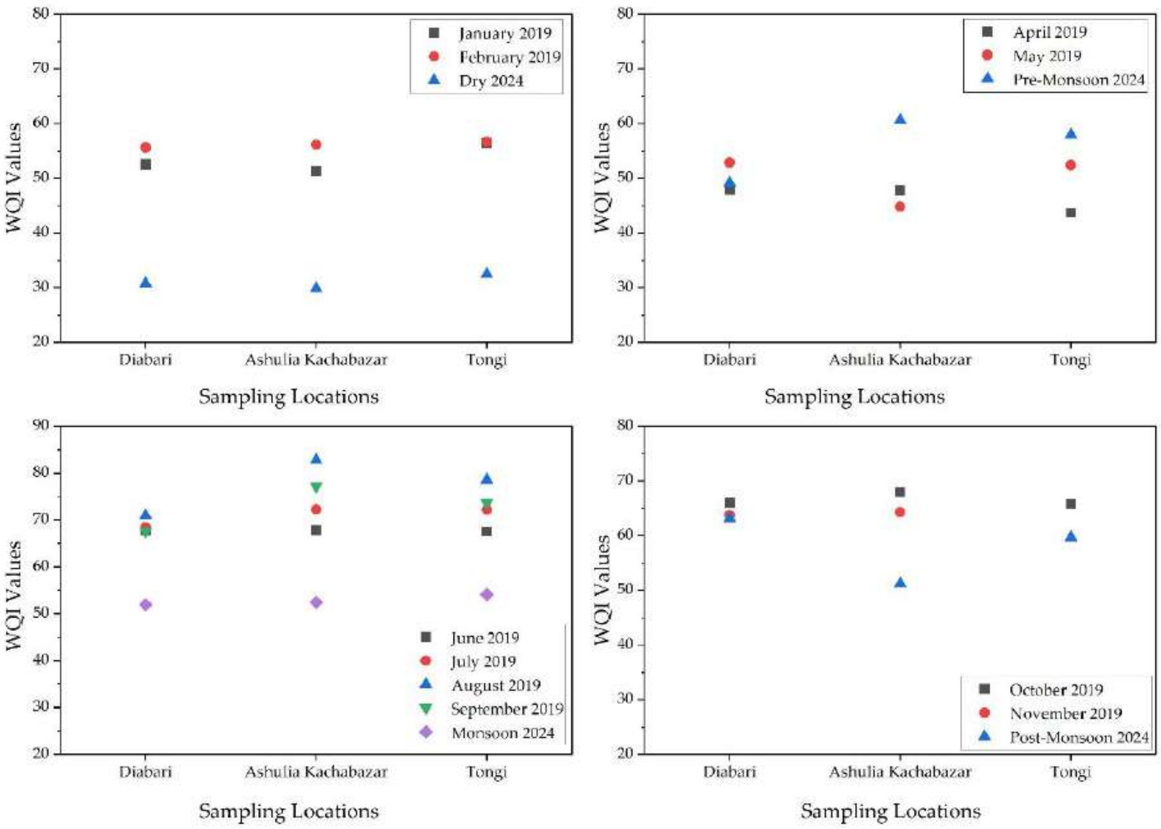


Figure 8.3: Comparison of WQI values across common sampling stations of Turag River

Figure 8.4 illustrates the WQI values for the Sitalakhya River. During both the dry and pre-monsoon seasons of 2019 and 2024, WQI values showed considerable variation across sampling stations. In the dry season, the South Rupshi station exhibited notably lower WQI values in 2024 compared to February 2019, whereas other stations recorded similar WQI values between the two years. During the pre-monsoon season, although WQI values were generally lower in 2024 than in 2019, the overall trend across stations remained consistent. In the monsoon season, WQI values were largely comparable between 2019 and 2024. However, during the post-monsoon season, the South Rupshi station experienced a decline in WQI values in 2024 relative to November 2019, while other stations showed higher WQI values in 2024 compared to the same period in 2019.

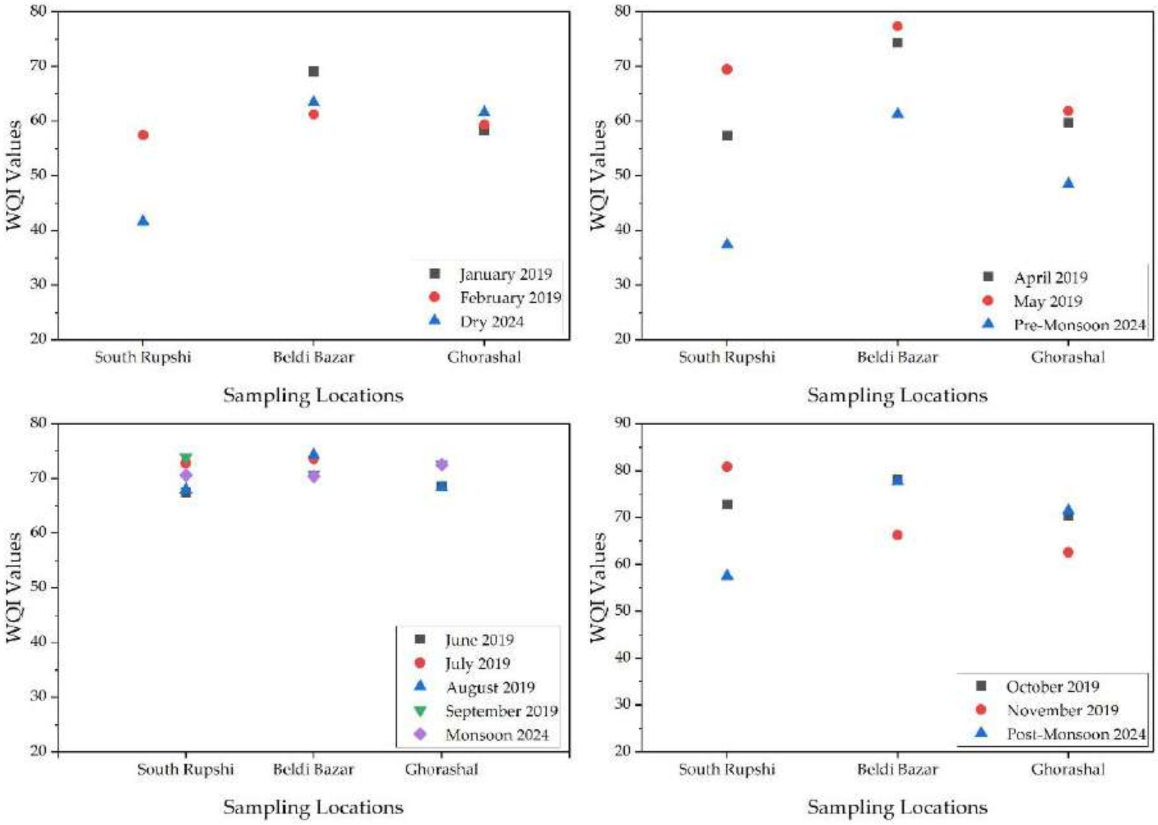


Figure 8.4: Comparison of WQI values across common sampling stations of Sitalakhya River

The comparative analysis of WQI values between 2019 and 2024 across the four rivers reveals a general decline in water quality over time, particularly evident during the dry and monsoon seasons. While some stations showed stable or only slight variations in water quality, others experienced significant deterioration, highlighting spatial variability in pollution and environmental stressors. The differences in seasonal trends and station-specific responses underscore the dynamic and complex nature of river water quality changes. These findings emphasize the need for continued and enhanced monitoring efforts to better understand temporal shifts and to guide effective water resource management strategies.

8.2 CCME WQI using the Eight Selected Parameters

The Canadian Council of Ministers of the Environment Water Quality Index (CCME WQI) is a widely recognized tool for assessing and communicating the overall quality of surface water. It synthesizes multiple water quality parameters into a single numerical value that reflects the degree to which a water body meets established water quality objectives or guidelines. The CCME WQI considers three factors: the scope (number of parameters that exceed guidelines), frequency (how often exceedances occur), and amplitude (the extent of exceedances), providing a comprehensive and interpretable measure of water quality. Due to its simplicity and robustness, the CCME WQI has been extensively applied in environmental monitoring programs globally, including river water quality assessments.

The CCME WQI offers a practical and efficient approach to assessing water quality, particularly due to its methodological simplicity and flexibility. Unlike traditional WQI models, it does not require the development of sub-indices or assignment of parameter-specific weightings, which are often subjective and time-consuming. This characteristic allows for a more rapid evaluation of water quality, making it suitable for timely decision-making and routine monitoring. Moreover, the CCME WQI imposes no restrictions on the number or type of parameters used, allowing it to accommodate a wide range of water quality indicators based on data availability and relevance to local conditions. This adaptability enhances the robustness of the index across diverse environmental contexts. As a result, the CCME WQI serves as a valuable tool for comprehensive and efficient water quality assessment, especially in cases where a standardized, scalable, and parameter-flexible approach is needed.

The CCME WQI was calculated seasonally for all four rivers to assess their water quality variations throughout the year. For each river, nine sampling stations were considered, with one observation from each station recorded per season, resulting in four CCME WQI values per river—one for each season. To ensure the assessment focused on the most critical factors affecting water quality, eight key parameters identified as most influential through Principal Component Analysis (PCA) were used for the CCME WQI calculation.

The CCME WQI values for the year 2024, summarized in Figure 8.5, reveal that all four rivers predominantly fall within the "Poor" category (WQI < 45) across most seasons, indicating frequent water quality guideline exceedances and overall degraded conditions. This consistent classification underscores widespread and sustained pollution pressure throughout the year.

In the Buriganga River, the CCME WQI values remained within the "Poor" category across all seasons, with a gradual increase from 26.46 in the dry season to 35.97 in the post-monsoon. While this upward trend suggests slight seasonal improvement, the overall water quality remains critically low. The Balu River followed a similar pattern, showing a modest seasonal improvement from 25.41 (dry) to 44.05 (post-monsoon), approaching the threshold for the "Marginal" category but still remaining classified as "Poor" in all seasons.

For the Turag River, the index showed the most notable seasonal variation. Starting with a low value of 25.85 in the dry season, the water quality improved considerably during the pre-monsoon (43.38) and reached 47.70 in the post-monsoon, the only instance among all rivers where a "Marginal" classification was achieved. This suggests a relative seasonal recovery in water quality, although levels still indicate frequent exceedances.

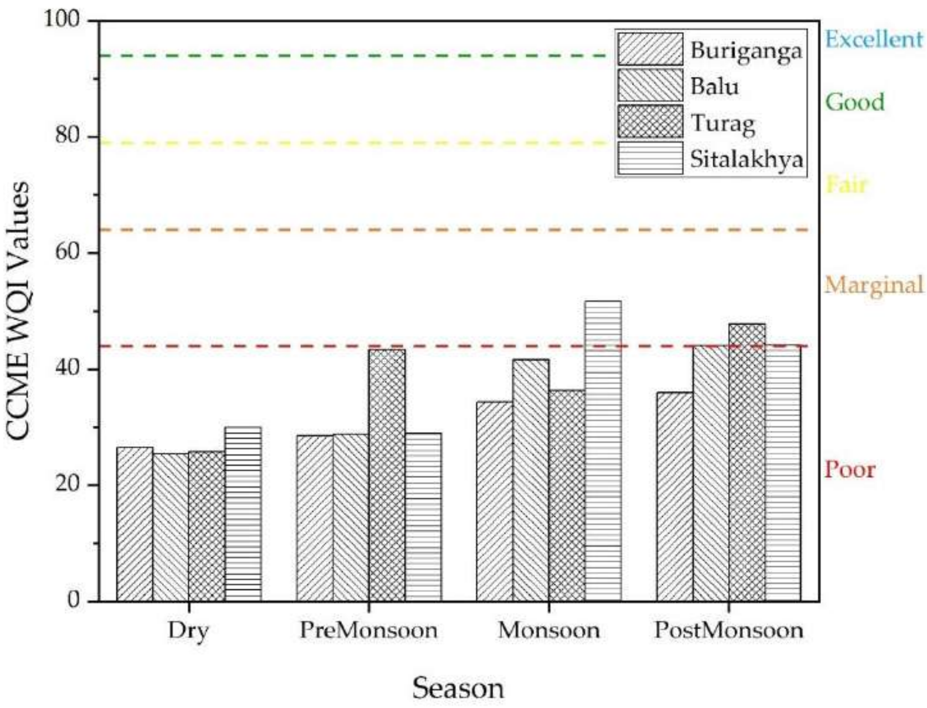


Figure 8.5: CCME WQI values for the four rivers based on 2024 data

The Sitalakhya River exhibited slightly better conditions overall. While CCME WQI values were still in the "Poor" range during the dry (29.99) and pre-monsoon (28.93) seasons, the river showed improved water quality during the monsoon season, with a WQI of 51.60, falling within the "Marginal" category. However, this improvement was not sustained, as values declined again to 44.18 in the post-monsoon, just below the marginal threshold.

Table 8.1: Seasonal CCME WQI distribution of the four rivers

WQI Value	Remarks	Count	Percentage (%)
95-100	Excellent	0	0
80-94	Good	0	0
65-79	Fair	0	0
45-64	Marginal	2	12.5
0-44	Poor	14	87.5
Total		16	100

Overall, the results highlight a seasonal trend of water quality improvement from the dry to monsoon seasons, likely due to increased dilution capacity during periods of higher flow. Nevertheless, most rivers failed to consistently rise above the "Poor" classification, emphasizing persistent water quality challenges. The Turag and Sitalakhya rivers demonstrated relatively better seasonal recovery, suggesting some variability in pollution sources or hydrological responses among the rivers. These findings correlate with the observations of our developed WQI model as well.

Chapter 9: Discussion and Strategic Recommendations

The findings from the seasonal and river-wise analysis of water quality reveal key trends and issues that need to be addressed for sustainable water resource management in and around Dhaka. Based on the observed WQI values, seasonal variations, and pollution sources, the following discussions are proposed to improve and monitor water quality more effectively:

1. Seasonal Variations in Water Quality and the Impact of Precipitation: Analyzing the WQI values for the four rivers across the different seasons, it becomes evident that water quality shows an overall improvement over time. Specifically, during the monsoon and post-monsoon seasons, the water quality of all rivers and sampling stations tends to improve when compared to the dry season. This seasonal variation is largely driven by changes in water flow, with the dry season seeing a notable decline in water quality. The reduced water flow during the dry season is significantly influenced by the conditions upstream of the rivers, leading to a higher concentration of pollutants. Various factors contribute to the decline in water quality during this period, such as industrial waste discharge, untreated sewage, agricultural runoff, and the disposal of solid waste. In contrast, the monsoon and post-monsoon seasons experience an influx of rainfall, which helps to dilute the pollutants and thus leads to an improvement in the water quality. This highlights the important role that precipitation plays in reducing the concentration of contaminants in the rivers.

2. Vertical Profiling in Water Quality Monitoring: In this study, water quality parameters were measured only from the surface layer of the river water. However, water quality can vary significantly at different depths due to factors such as stratification, sediment interactions, and varying levels of dissolved oxygen or pollutants. Relying solely on surface water data may not fully represent the overall water quality conditions of the river. Therefore, future studies are recommended to incorporate sampling from multiple water depths. This approach would provide a more complete and accurate understanding of vertical variations in water quality, leading to more informed and effective river management strategies.

3. Regular and Real-Time Water Quality Monitoring: To enhance the accuracy and reliability of water quality assessments, it is essential to implement regular and real-time monitoring systems. In this study, only one observation per season was used at each sampling station, which is not sufficient to fully capture the temporal variability and dynamic nature of river pollution. Integrating continuous or more frequent monitoring—through spectrophotometry and other established methods—can significantly improve the quality and depth of the data collected. Real-time monitoring, in particular, allows for the immediate detection of pollution events and supports the timely calculation of the WQI. This approach strengthens the WQI’s role as a responsive and informative tool for water management, enabling quicker interventions and a more comprehensive understanding of long-term water quality trends.

4. Emerging Pollutants: Toxic, Carcinogenic, and Antibiotic Contaminants: The current WQI framework used in this study does not account for emerging pollutants such as toxic chemicals, carcinogenic compounds, and antibiotic residues, which are increasingly present in urban river systems. These pollutants, though often found in low concentrations, can pose serious risks to aquatic life, particularly to primary producers and fish populations. Their long-term presence can disrupt aquatic ecosystems, bioaccumulate through food chains, and even threaten human health. Future monitoring and WQI development should incorporate these contaminants to better assess ecological risks and determine whether aquatic organisms can survive and thrive under current water quality conditions. Including such parameters will enhance the ecological relevance and predictive power of the WQI for sustainable river management.

5. Targeted Pollution Management to Address Seasonal Fluctuations and River-Specific Variability: The WQI values reveal that the water quality of the rivers fluctuates between the "bad" and "good" categories throughout the year, with the most noticeable deterioration occurring during the dry season. During this time, the water quality tends to degrade more significantly due to the reduced flow and the increased pollution load. However, when rainfall occurs in the monsoon and post-monsoon seasons, there is a marked improvement in water quality, primarily because the rainfall dilutes the pollutants that accumulate in the rivers. This demonstrates the positive effect of rainfall in mitigating water quality deterioration. Among the rivers studied, the Buriganga River shows relatively minimal variation in water quality over time, remaining within a similar range across all seasons. On the other hand, the Balu and Sitalakhya rivers show more substantial fluctuations in water quality, with more pronounced variations observed between the dry and monsoon/post-monsoon seasons. Despite the heavy pollution load experienced during the dry season, the water quality of all rivers remains within the "bad" category and does not fall into the "very bad" category. This suggests that even under conditions of high pollution, the water quality does not reach an extremely poor state. It also indicates that with the right interventions, such as better pollution management and regulatory measures, it is possible to prevent further deterioration and even improve water quality over time.

6. Strengthening Pollution Control and Legislative Measures for Sustainable Water Quality Improvement: During the monsoon and post-monsoon seasons, the water quality of the rivers generally falls within the "medium" category, with some areas achieving a "good" water quality status. The rainfall during these seasons helps to reduce the pollutant load, leading to improved water conditions. However, the water quality could be further enhanced with a more significant reduction in pollution levels. In these seasons, while the improvement is noticeable, further efforts are necessary to achieve even better-quality levels. Long-term improvements in water quality will require the implementation of comprehensive legislative measures that aim to reduce pollution. Effective enforcement of these measures, coupled with efficient water management strategies, will be essential in maintaining and further improving the water quality, particularly during the challenging dry season when water quality is most vulnerable to degradation. Through such actions, it is possible to stabilize and enhance water quality across all seasons, ensuring healthier river ecosystems in the long run.

7. River-Specific Statistical Analyses for Targeted Pollution Management: Using river-specific statistical tools can help identify the most influential pollution indicators unique to each river system. Techniques such as multivariate analysis, clustering, or dimensionality reduction methods can be applied separately to each river to isolate the key factors affecting water quality. This approach considers local conditions, pollution sources, and seasonal variations, enabling the identification of critical parameters that influence river health. To ensure the reliability of these analyses, it is important to collect comprehensive, high-frequency water quality data. Such river-wise statistical assessments support the development of targeted pollution control strategies tailored to the unique pollution dynamics of each river, enhancing the effectiveness of monitoring and leading to more informed, site-specific decision-making for conservation and management efforts.

Based on the analytical findings, spatial-temporal patterns, and methodological insights presented in this study, several key areas for improvement in river water quality monitoring and management have been identified. The observed limitations in data resolution, parameter coverage, and pollution source differentiation suggest that enhancements in both the technical and institutional frameworks are needed. The following recommendations are proposed to strengthen the accuracy, applicability, and policy relevance of the Water Quality Index (WQI) and to support more informed decision-making for sustainable river basin management.

- Conduct more frequent and real-time monitoring to capture short-term pollution fluctuations and improve WQI responsiveness.
- Include emerging pollutants like heavy metals, microplastics, and antibiotics to better assess ecological and health risks.
- Incorporate depth-based sampling to capture vertical variations in water quality.
- Use spatial and statistical tools to distinguish pollution sources for targeted intervention.
- Revise sub-indexing and weighting to reflect diverse water uses and parameter relevance.
- Combine PCA with other statistical methods to strengthen parameter selection.
- Apply GIS tools for pollution mapping and water quality management.

WQI is a practical tool for continuous monitoring and proactive water quality management. Despite its limitations, the newly developed WQI model serves as a valuable and effective tool for water quality management, providing a systematic approach for monitoring water quality over time. By regularly assessing water quality using the WQI, authorities can track changes and trends, identifying potential areas of concern before they worsen. The WQI's role in continuous monitoring allows for the early detection of sudden deteriorations in water quality, which may be caused by pollution events such as industrial discharges, waste dumping, or seasonal changes. When a decline in water quality is detected, immediate action can be taken to address the issue before it leads to irreversible damage. Moreover, the WQI serves as an important indicator for pinpointing specific areas where pollution levels are on the rise, thereby enabling targeted and timely interventions. This helps prevent further degradation of water quality, particularly in areas most affected by pollution sources. By identifying pollution hotspots, decision-makers can prioritize regions that require urgent attention, ensuring that the appropriate measures are put in place to reduce pollution levels effectively. The WQI's role extends beyond tracking water quality—it also acts as a catalyst for action, promoting proactive solutions for water management. The developed WQI model can help in the water quality monitoring system by the following ways:

- **Continuous Water Quality Monitoring**: The WQI allows for regular and continuous tracking of water quality over time, providing real-time data on changes in water conditions.
- **Identification of Pollution Hotspots**: By analyzing WQI data, areas with rising pollution levels can be identified, helping to pinpoint specific locations that require urgent intervention.
- **Assessment of Seasonal Fluctuations**: The WQI accounts for seasonal variations in water quality, aiding in understanding the impacts of different seasons (dry, monsoon, and post-monsoon) on water conditions.
- **Tracking Long-Term Water Quality Trends**: The model helps in monitoring long-term water quality trends, providing insights into improvements or deteriorations over extended periods.
- **Enhancing Public Awareness**: By providing a clear, understandable index of water quality, the WQI helps raise public awareness about the state of water resources and the need for conservation.
- **Guiding River Ecosystem Protection**: The WQI plays a crucial role in maintaining healthy river ecosystems by facilitating the identification and mitigation of factors that threaten water quality.
- **Improved Water Quality Management Framework**: By combining data monitoring with interventions, the WQI offers a structured framework for managing water quality in the long term.

Chapter 10: Conclusion

This study presents a structured and data-driven approach to assess the surface water quality of four major rivers in the Dhaka region, Buriganga, Turag, Balu, and Sitalakhya, through the development of a Water Quality Index (WQI) model. The model was established using Principal Component Analysis (PCA), which allowed the selection of the most influential water quality parameters based on statistical correlation and variance. By integrating multiple parameters into a single index, the WQI provides an efficient and interpretable means to assess and compare river water quality across different locations and seasons.

The data collection for this study was carefully designed to capture seasonal and spatial variations in water quality. Sampling was conducted throughout the year 2024, covering all four major seasons to reflect changing environmental conditions. At each of the four rivers nine fixed sampling points were selected to represent different stretches and pollution gradients. The data collection process combined both onsite measurements and laboratory analyses, ensuring accurate and comprehensive assessment of key water quality parameters. This extensive and systematic sampling approach provided a robust dataset that underpins the development and validation of the Water Quality Index model presented in this study.

The findings reveal clear seasonal variation in water quality. The most degraded conditions were observed during the dry season, likely due to low river flow, higher pollutant concentration, and limited dilution capacity. In contrast, the monsoon and post-monsoon seasons showed moderate improvement, largely as a result of increased rainfall and flow that helped flush pollutants downstream. Among the four rivers studied, the Buriganga consistently demonstrated the poorest water quality throughout the year, confirming its status as the most polluted among the group. Meanwhile, the upper section of the Sitalakhya River generally exhibited better water quality, particularly during the post-monsoon period.

In addition to the WQI calculation, the study incorporated several analytical tools to enhance the reliability and communication of results. ArcGIS-based spatial visualization was used to map WQI values along the river stretches, effectively identifying pollution hotspots and spatial patterns. The inclusion of stratified bootstrap analysis added a layer of uncertainty evaluation, offering confidence intervals around WQI estimates and highlighting data variability. These components together made the assessment more comprehensive and transparent.

The application of the Canadian Council of Ministers of the Environment Water Quality Index (CCME WQI) to the 2024 dataset provided additional validation. It confirmed that water quality in all four rivers generally remains in the "Poor" category, with only marginal seasonal improvements. These consistent patterns suggest that natural seasonal recovery is insufficient to overcome the ongoing pollution loads, particularly in urban and downstream areas.

Despite its strengths, the study has several limitations. All water quality samples were collected from the surface layer of the rivers, which does not reflect potential vertical variation. Only one observation per season per station was used, limiting the ability to detect short-term or event-based pollution changes. Moreover, the WQI model did not incorporate emerging pollutants such as heavy metals, microplastics, or antibiotic residues, which are increasingly relevant in urban river systems. The model also does not distinguish between pollution sources, which is essential for targeted intervention.

To improve the accuracy and applicability of the WQI in future studies, several enhancements are recommended. These include the implementation of real-time and more frequent monitoring, inclusion of depth-stratified sampling, expansion of the parameter list to include emerging contaminants, and use of statistical tools to differentiate pollution sources. Refinement of sub-indexing and parameter weighting methods to reflect multiple water uses, such as irrigation, fisheries, and domestic supply, would also increase the index's relevance to policy and regulatory frameworks.

Overall, this study contributes a practical and adaptable methodology for evaluating the water quality of rivers in highly urbanized and industrialized regions. The developed WQI model, supported by PCA, CCME assessment, spatial mapping, and uncertainty analysis, offers a replicable framework for future river basin monitoring in Bangladesh. It can serve as a decision-support tool for environmental agencies, urban planners, and policymakers working toward sustainable river management. With further refinement and institutional support, the WQI model can play an important role in guiding water quality protection and restoration efforts in the region.

Chapter 11: References

2010 ENVIRONMENTAL PERFORMANCE INDEX. (n.d.).

Abbasi, T., & Abbasi, S. A. (2012a). Water Quality Indices. *Water Quality Indices*, 1–362. https://doi.org/10.1016/C2010-0-69472-7

Abbasi, T., & Abbasi, S. A. (2012b). Water-Quality Indices. In *Water Quality Indices* (pp. 353–356). Elsevier. https://doi.org/10.1016/B978-0-444-54304-2.00016-6

Abbasi, T., & Abbasi, S. A. (2012c). Water-Quality Indices: Looking Back, Looking Ahead. *Water Quality Indices*, 353–356. https://doi.org/10.1016/B978-0-444-54304-2.00016-6

Abbasi, T., & Abbasi, S. A. (2012d). Water-Quality Indices of USA and Canada. *Water Quality Indices*, 175–186. https://doi.org/10.1016/B978-0-444-54304-2.00010-5

Abrahão, R., Carvalho, M., Da Silva Jr, W., Machado, T., Gadelha, C., & Hernandez, M. (2010). Use of index analysis to evaluate the water quality of a stream receiving industrial effluents. *Water SA*, *33*(4). https://doi.org/10.4314/wsa.v33i4.52940

Abrahão, R., Carvalho, M., Da Silva, W. R., Machado, T. T. V., Gadelha, C. L. M., & Hernandez, M. I. M. (2007). Use of index analysis to evaluate the water quality of a stream receiving industrial effluents. *Water SA*, *33*(4), 459–465. https://doi.org/10.4314/WSA.V33I4.52940

Acquavita, A., Aleffi, I. F., Benci, C., Bettoso, N., Crevatin, E., Milani, L., Tamberlich, F., Toniatti, L., Barbieri, P., Licen, S., & Mattassi, G. (2015). Annual characterization of the nutrients and trophic state in a Mediterranean coastal lagoon: The Marano and Grado Lagoon (northern Adriatic Sea). *Regional Studies in Marine Science*, *2*, 132–144. https://doi.org/10.1016/J.RSMA.2015.08.017

Adams, S., Titus, R., Pietersen, K., Tredoux, G., & Harris, C. (2001). Hydrochemical characteristics of aquifers near Sutherland in the Western Karoo, South Africa. *Journal of Hydrology*, *241*(1–2), 91–103. https://doi.org/10.1016/S0022-1694(00)00370-X

Akay, E., & Dalkıran, N. (2020). Assessing biological water quality of Yalakdere stream (Yalova, Turkey) with benthic macroinvertebrate-based metrics. *Biologia*, *75*(9), 1347–1363. https://doi.org/10.2478/S11756-019-00387-9/METRICS

Akbor, Md. A., Uddin, M., & Ahsan, Md. A. (2017). Investigation of Water Quality Parameters at Different Points in the Buriganga River, Bangladesh. *Journal of Environmental Science and Natural Resources*, *10*(1), 75–80. https://doi.org/10.3329/JESNR.V10I1.34698

Al-Afify, A. D. G., Tahoun, U. M., & Abdo, M. H. (2019). Water Quality Index and Microbial Assessment of Lake Qarun, El-Batts and El-Wadi Drains, Fayoum Province, Egypt. *Egyptian Journal of Aquatic Biology and Fisheries*, *23*(1), 341–357. https://doi.org/10.21608/EJABF.2019.28270

Ali, S. A., & Ali, U. (2013). Litho-Structural Mapping of Sind Catchment (Kashmir Basin), NW Himalaya, using Remote Sensing & GIS Techniques. *International Journal of Science and Research*, *4*, 2319–7064.

Ali, S. A., & Ali, U. (2018). Hydrochemical characteristics and spatial analysis of groundwater quality in parts of Bundelkhand Massif, India. *Applied Water Science*, *8*(1), 1–15. https://doi.org/10.1007/S13201-018-0678-X/FIGURES/10

Ali, S. A., & Pirasteh, S. (2004). Geological applications of Landsat Enhanced Thematic Mapper (ETM) data and Geographic Information System (GIS): Mapping and structural interpretation in south-west Iran, Zagros Structural Belt. *International Journal of Remote Sensing*, *25*(21), 4715–4727. https://doi.org/10.1080/01431160410001688295

Almeida, C., González, S. O., Mallea, M., & González, P. (2012). A recreational water quality index using chemical, physical and microbiological parameters. *Environmental Science and Pollution Research*, *19*(8), 3400–3411. https://doi.org/10.1007/S11356-012-0865-5/FIGURES/8

Anderson, D. M., Fensin, E., Gobler, C. J., Hoeglund, A. E., Hubbard, K. A., Kulis, D. M., Landsberg, J. H., Lefebvre, K. A., Provoost, P., Richlen, M. L., Smith, J. L., Solow, A. R., & Trainer, V. L. (2021). Marine harmful algal blooms (HABs) in the United States:

History, current status and future trends. *Harmful Algae*, *102*, 101975. https://doi.org/10.1016/J.HAL.2021.101975

Babaei Semiromi, F., Hassani, A. H., Torabian, A., Karbassi, A. R., & Hosseinzadeh Lotfi, F. (2011). Evolution of a new surface water quality index for Karoon catchment in Iran. *Water Science and Technology*, *64*(12), 2483–2491. https://doi.org/10.2166/WST.2011.780

Batur, E., & Maktav, D. (2019). Assessment of Surface Water Quality by Using Satellite Images Fusion Based on PCA Method in the Lake Gala, Turkey. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(5), 2983–2989. https://doi.org/10.1109/TGRS.2018.2879024

Belle-Oudry, D. (2008). Quantitative Analysis of Sulfate in Water by Indirect EDTA Titration. *Journal of Chemical Education*, *85*(9), 1269–1270. https://doi.org/10.1021/ED085P1269

Bengraïne, K., & Marhaba, T. F. (2003). Using principal component analysis to monitor spatial and temporal changes in water quality. *Journal of Hazardous Materials*, *100*(1–3), 179–195. https://doi.org/10.1016/S0304-3894(03)00104-3

Benkov, I., Varbanov, M., Venelinov, T., & Tsakovski, S. (2023). Principal Component Analysis and the Water Quality Index—A Powerful Tool for Surface Water Quality Assessment: A Case Study on Struma River Catchment, Bulgaria. *Water 2023, Vol. 15, Page 1961*, *15*(10), 1961. https://doi.org/10.3390/W15101961

Bigham, T., Dooley, J. S. G., Ternan, N. G., Snelling, W. J., Héctor Castelán, M. C., & Davis, J. (2019). Assessing microbial water quality: Electroanalytical approaches to the detection of coliforms. *TrAC Trends in Analytical Chemistry*, *121*, 115670. https://doi.org/10.1016/J.TRAC.2019.115670

Bonadonna, L., Briancesco, R., & La Rosa, G. (2019). Innovative analytical methods for monitoring microbiological and virological water quality. *Microchemical Journal*, *150*, 104160. https://doi.org/10.1016/J.MICROC.2019.104160

Bordalo, A. A., Teixeira, R., & Wiebe, W. J. (2006). A water quality index applied to an international shared river basin: The case of the Douro River. *Environmental*

Management, *38*(6), 910–920. https://doi.org/10.1007/S00267-004-0037-6/FIGURES/7

Bozorg-Haddad, O., Delpasand, M., & Loáiciga, H. A. (2021). Water quality, hygiene, and health. *Economical, Political, and Social Issues in Water Resources*, 217–257. https://doi.org/10.1016/B978-0-323-90567-1.00008-5

Brown, R. M., McClelland, N. I., Deininger, R. A., & Landwehr, J. M. (1973). Validating the WQI. *Proceedings of the National Meeting of American Society of Civil Engineers on Water Resources Engineering, Washington, DC, USA*, *29*.

Bu, J., Liu, W., Pan, Z., & Ling, K. (2020). Comparative Study of Hydrochemical Classification Based on Different Hierarchical Cluster Analysis Methods. *International Journal of Environmental Research and Public Health 2020, Vol. 17, Page 9515*, *17*(24), 9515. https://doi.org/10.3390/IJERPH17249515

BWDB. (2023). *Flood Forecasting and Warning Centre, BWDB, Bangladesh*. http://www.ffwc.gov.bd/

Carvalho, L., Cortes, R., & Bordalo, A. A. (2011). Evaluation of the ecological status of an impaired watershed by using a multi-index approach. *Environmental Monitoring and Assessment*, *174*(1–4), 493–508. https://doi.org/10.1007/S10661-010-1473-9/METRICS

CCME WATER QUALITY INDEX USER'S MANUAL 2017 UPDATE. (2017).

Chan, J. Y. Le, Leow, S. M. H., Bea, K. T., Cheng, W. K., Phoong, S. W., Hong, Z. W., & Chen, Y. L. (2022). Mitigating the Multicollinearity Problem and Its Machine Learning Approach: A Review. *Mathematics 2022, Vol. 10, Page 1283*, *10*(8), 1283. https://doi.org/10.3390/MATH10081283

Chauhan, V. S., & Sharma, A. (2003). Studies on organoleptic properties of food products from fresh egg and egg powder through principal component analysis. *Food / Nahrung*, *47*(2), 102–105. https://doi.org/10.1002/FOOD.200390018

Chowdhary, P., Bharagava, R. N., Mishra, S., Khan, N., Chowdhary, P., Bharagava, · R N, Mishra, S., & Khan, · N. (2020). Role of Industries in Water Scarcity and Its Adverse

Effects on Environment and Human Health. *Environmental Concerns and Sustainable Development*, 235–256. https://doi.org/10.1007/978-981-13-5889-0_12

Cooper, C. M. (1993). Biological Effects of Agriculturally Derived Surface Water Pollutants on Aquatic Systems—A Review. *Journal of Environmental Quality*, *22*(3), 402–408. https://doi.org/10.2134/JEQ1993.00472425002200030003X

Country Profile Bangladesh COUNTRY OVERVIEW. (2011).

Country profile-Bangladesh. (2014).

Cude, C. G. (2001). OREGON WATER QUALITY INDEX A TOOL FOR EVALUATING WATER QUALITY MANAGEMENT EFFECTIVENESS1. *JAWRA Journal of the American Water Resources Association*, *37*(1), 125–137. https://doi.org/10.1111/J.1752-1688.2001.TB05480.X

Dadolahi-Sohrab, A., Arjomand, F., & Fadaei-Nasab, M. (2012). Water quality index as a simple indicator of watersheds pollution in southwestern part of Iran. *Water and Environment Journal*, *26*(4), 445–454. https://doi.org/10.1111/J.1747-6593.2011.00303.X

Devendra Swaroop Bhargave, B., & Asce, F. (1985). Expression for Drinking Water Supply Standards. *Journal of Environmental Engineering*, *111*(3), 304–316. https://doi.org/10.1061/(ASCE)0733-9372(1985)111:3(304)

Dhany Sutadian, A., Muttil, N., Yilmaz, A., & Perera, C. (n.d.). *Development of River Water Quality Indices-A Review*.

DoE. (2024, July 27). *SURFACE AND GROUND WATER QUALITY REPORT 2023*. https://doe.portal.gov.bd/sites/default/files/files/doe.portal.gov.bd/publications/edc0f4f2_8346_4f58_bdd1_72978984ba3e/2024-08-27-08-00-a698975451058b1fc83366d454942747.pdf

Dojlido, J., Raniszewski, J., & Woyciechowska, J. (1994). Water quality index applied to rivers in the Vistula river basin in Poland. *Environmental Monitoring and Assessment*, *33*(1), 33–42. https://doi.org/10.1007/BF00546659/METRICS

Donald, A. E., Blessing, U. A., Anyanwu, C., & Donald, E. (2019). Index approach to water quality assessment of a south eastern Nigerian river. *~ 153 ~ International Journal of Fisheries and Aquatic Studies*, *7*(1).

DPHE. (2019, May 15). *জনস্বাস্থ্য প্রকৌশল অধিদপ্তর*. https://dphe.gov.bd/site/page/15fa0d7b-11f1-45c0-a684-10a543376873/Water-Quality-Parameters-

Duclos Alegue, J., & Gnauck, A. (2006). *Signal Analysis of Water Quality Indicators*.

ECR. (2023, May 3). *Bangladesh Environmental Conservation Rules 2023*. https://file-chittagong.portal.gov.bd/uploads/c2caafaa-e471-422d-9344-8bbc195d1a74/641/180/9a8/6411809a82e0b221760915.pdf

Effendi, H., Romanto, & Wardiatno, Y. (2015a). Water Quality Status of Ciambulawung River, Banten Province, Based on Pollution Index and NSF-WQI. *Procedia Environmental Sciences*, *24*, 228–237. https://doi.org/10.1016/J.PROENV.2015.03.030

Effendi, H., Romanto, & Wardiatno, Y. (2015b). Water Quality Status of Ciambulawung River, Banten Province, Based on Pollution Index and NSF-WQI. *Procedia Environmental Sciences*, *24*, 228–237. https://doi.org/10.1016/j.proenv.2015.03.030

Ewaid, S. H. (2017). Water quality evaluation of Al-Gharraf river by two water quality indices. *Applied Water Science*, *7*(7), 3759–3765. https://doi.org/10.1007/S13201-016-0523-Z/TABLES/7

FAO. (2023). *Water quality in agriculture: Risks and risk mitigation*. FAO; IWMI; https://doi.org/10.4060/cc7340en

Fartas, F., Remini, B., Sekiou, F., & Marouf, N. (2022). The use of PCA and ANN to improve evaluation of the WQIclassic, development of a new index, and prediction of WQI, Coastel Constantinois, northern coast of eastern Algeria. *Water Supply*, *22*(12), 8727–8749. https://doi.org/10.2166/ws.2022.389

Fatema, K., Begum, M., Zahid, M. Al, & Hossain, M. (2018). Water quality assessment of the river Buriganga, Bangladesh. *Journal of Biodiversity Conservation and Bioresource Management*, *4*(1), 47–54. https://doi.org/10.3329/JBCBM.V4I1.37876

Filik Iscen, C., Emiroglu, Ö., Ilhan, S., Arslan, N., Yilmaz, V., & Ahiska, S. (2008). Application of multivariate statistical techniques in the assessment of surface water quality in Uluabat Lake, Turkey. *Environmental Monitoring and Assessment*, *144*(1–3), 269–276. https://doi.org/10.1007/S10661-007-9989-3/METRICS

Garg, S., Chowdhury, Z. Z., Faisal, A. N. M., Rumjit, N. P., & Thomas, P. (2022). Impact of Industrial Wastewater on Environment and Human Health. *Environmental Science and Engineering*, 197–209. https://doi.org/10.1007/978-3-030-83811-9_10/TABLES/1

Gazzaz, N. M., Yusoff, M. K., Aris, A. Z., Juahir, H., & Ramli, M. F. (2012). Artificial neural network modeling of the water quality index for Kinta River (Malaysia) using water quality variables as predictors. *Marine Pollution Bulletin*, *64*(11), 2409–2420. https://doi.org/10.1016/J.MARPOLBUL.2012.08.005

Gholizadeh, M., Melesse, A., & Reddi, L. (2016). A Comprehensive Review on Water Quality Parameters Estimation Using Remote Sensing Techniques. *Sensors*, *16*(8), 1298. https://doi.org/10.3390/s16081298

Girones, R., Ferrús, M. A., Alonso, J. L., Rodriguez-Manzano, J., Calgua, B., de Abreu Corrêa, A., Hundesa, A., Carratala, A., & Bofill-Mas, S. (2010). Molecular detection of pathogens in water – The pros and cons of molecular techniques. *Water Research*, *44*(15), 4325–4339. https://doi.org/10.1016/J.WATRES.2010.06.030

Goodman, C. M. (1987). The Delphi technique: A critique. *Journal of Advanced Nursing*, *12*(6), 729–734. https://doi.org/10.1111/J.1365-2648.1987.TB01376.X

Gupta, S., & Gupta, S. K. (2021). A critical review on water quality index tool: Genesis, evolution and future directions. *Ecological Informatics*, *63*, 101299. https://doi.org/10.1016/J.ECOINF.2021.101299

Hasan, H. H., Jamil, N. R., & Aini, N. (2015). Water Quality Index and Sediment Loading Analysis in Pelus River, Perak, Malaysia. *Procedia Environmental Sciences*, *30*, 133–138. https://doi.org/10.1016/J.PROENV.2015.10.024

Horton, R. K. (1965). An index number system for rating water quality. *J Water Pollut Control Fed*, *37*(3), 300–306.

Hoseinzadeh, E., Khorsandi, H., Wei, C., & Alipour, M. (2015). Evaluation of Aydughmush River water quality using the National Sanitation Foundation Water Quality Index (NSFWQI), River Pollution Index (RPI), and Forestry Water Quality Index (FWQI). *Desalination and Water Treatment*, *54*(11), 2994–3002. https://doi.org/10.1080/19443994.2014.913206

House, M. A. (1989). A Water Quality Index for River Management. *Water and Environment Journal*, *3*(4), 336–344. https://doi.org/10.1111/j.1747-6593.1989.tb01538.x

Hsu, C.-C., & Sandford, B. A. (2019). The Delphi Technique: Making Sense of Consensus. *Practical Assessment, Research, and Evaluation*, *12*(1), 10. https://doi.org/10.7275/pdz9-th90

Huang, F., Wang, X., Lou, L., Zhou, Z., & Wu, J. (2010). Spatial variation and source apportionment of water pollution in Qiantang River (China) using statistical techniques. *Water Research*, *44*(5), 1562–1572. https://doi.org/10.1016/J.WATRES.2009.11.003

Hudson, N., Baker, A., Ward, D., Reynolds, D. M., Brunsdon, C., Carliell-Marquet, C., & Browning, S. (2008). Can fluorescence spectrometry be used as a surrogate for the Biochemical Oxygen Demand (BOD) test in water quality assessment? An example from South West England. *Science of The Total Environment*, *391*(1), 149–158. https://doi.org/10.1016/J.SCITOTENV.2007.10.054

Huey, G. M., & Meyer, M. L. (2010). Turbidity as an Indicator of Water Quality in Diverse Watersheds of the Upper Pecos River Basin. *Water 2010, Vol. 2, Pages 273-284*, *2*(2), 273–284. https://doi.org/10.3390/W2020273

IONUŞ, O. (2010). WATER QUALITY INDEX-ASSESSMENT METHOD OF THE MOTRU RIVER WATER QUALITY (OLTENIA, ROMANIA) INDICELE DE CALITATE A APEI–METODĂ DE DETERMINARE A CALITĂ\c{T}II APEI RÂULUI MOTRU (OLTENIA, ROMÂNIA). *OF THE UNIVERSITY OF CRAIOVA*, 74.

Jollife, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical,* *Physical and Engineering Sciences*, *374*(2065). https://doi.org/10.1098/RSTA.2015.0202

Juahir, H., Zain, S. M., Yusoff, M. K., Hanidza, T. I. T., Armi, A. S. M., Toriman, M. E., & Mokhtar, M. (2011). Spatial water quality assessment of Langat River Basin (Malaysia) using environmetric techniques. *Environmental Monitoring and Assessment*, *173*(1–4), 625–641. https://doi.org/10.1007/S10661-010-1411-X/METRICS

Kabir, A., Sraboni, H. J., Hasan, M. M., & Sorker, R. (2022). Eco-environmental assessment of the Turag River in the megacity of Bangladesh. *Environmental Challenges*, *6*, 100423. https://doi.org/10.1016/J.ENVC.2021.100423

Kamal, M. M., Malmgren-Hansen, A., & Badruzzaman, A. B. M. (1999). Assessment of pollution of the river Buriganga, Bangladesh, using a water quality model. *Water Science and Technology*, *40*(2), 129–136. https://doi.org/10.2166/WST.1999.0104

Kannel, P. R., Lee, S., Lee, Y. S., Kanel, S. R., & Khan, S. P. (2007). Application of water quality indices and dissolved oxygen as indicators for river water classification and urban impact assessment. *Environmental Monitoring and Assessment*, *132*(1–3), 93–110. https://doi.org/10.1007/S10661-006-9505-1/METRICS

Kasiske, D., Klinkmüller, K. D., & Sonneborn, M. (1978). Application of high-performance liquid chromatography to water pollution analysis. *Journal of Chromatography A*, *149*(C), 703–710. https://doi.org/10.1016/S0021-9673(00)81022-5

Kazi, T. G., Arain, M. B., Jamali, M. K., Jalbani, N., Afridi, H. I., Sarfraz, R. A., Baig, J. A., & Shah, A. Q. (2009). Assessment of water quality of polluted lake using multivariate statistical techniques: A case study. *Ecotoxicology and Environmental Safety*, *72*(2), 301–309. https://doi.org/10.1016/J.ECOENV.2008.02.024

Khan, M. H., Akash, N. M., Akter, S., Rukh, M., Nzediegwu, C., & Islam, M. S. (2023). A comprehensive review of coconut-based porous materials for wastewater treatment and CO2 capture. *Journal of Environmental Management*, *338*, 117825. https://doi.org/10.1016/J.JENVMAN.2023.117825

Khatun, M. A., Rashid, M. B., & Hygen, H. O. (n.d.). *Climate of Bangladesh*.

Khuan, L. Y., Hamzah, N., & Jailani, R. (2002). Prediction of water quality index (WQI) based on artificial neural network (ANN). *2002 Student Conference on Research and Development: Globalizing Research and Development in Electrical and Electronics Engineering, SCOReD 2002 - Proceedings*, 157–161. https://doi.org/10.1109/SCORED.2002.1033081

Koçer, M. A. T., & Sevgili, H. (2014). Parameters selection for water quality index in the assessment of the environmental impacts of land-based trout farms. *Ecological Indicators*, *36*, 672–681. https://doi.org/10.1016/J.ECOLIND.2013.09.034

Lermontov, A., Yokoyama, L., Lermontov, M., & Machado, M. A. S. (2009). River quality analysis using fuzzy water quality index: Ribeira do Iguape river watershed, Brazil. *Ecological Indicators*, *9*(6), 1188–1197. https://doi.org/10.1016/J.ECOLIND.2009.02.006

Li, J., Abdulmohsin, H. A., Hasan, S. S., Kaiming, L., Al-Khateeb, B., Ghareb, M. I., & Mohammed, M. N. (2019). Hybrid soft computing approach for determining water quality indicator: Euphrates River. *Neural Computing and Applications*, *31*(3), 827–837. https://doi.org/10.1007/S00521-017-3112-7/TABLES/2

Linstone, H. A. (1985). The Delphi technique. *Environmental Impact Assessment, Technology Assessment, and Risk Analysis. Proc., Les Arcs, 1983*, 621–649. https://doi.org/10.1007/978-3-642-70634-9_22/COVER

Liou, S. M., Lo, S. L., & Wang, S. H. (2004). A generalized water quality index for Taiwan. *Environmental Monitoring and Assessment*, *96*(1–3), 35–52. https://doi.org/10.1023/B:EMAS.0000031715.83752.A1/METRICS

Liou, S.-M., Lo, S.-L., & Wang, S.-H. (2004). A Generalized Water Quality Index for Taiwan. *Environmental Monitoring and Assessment*, *96*(1–3), 35–52. https://doi.org/10.1023/B:EMAS.0000031715.83752.a1

Liu, C. W., Lin, K. H., & Kuo, Y. M. (2003). Application of factor analysis in the assessment of groundwater quality in a blackfoot disease area in Taiwan. *Science of The Total Environment*, *313*(1–3), 77–89. https://doi.org/10.1016/S0048-9697(02)00683-6

Lobato, T. C., Hauser-Davis, R. A., Oliveira, T. F., Silveira, A. M., Silva, H. A. N., Tavares, M. R. M., & Saraiva, A. C. F. (2015a). Construction of a novel water quality index and quality indicator for reservoir water quality evaluation: A case study in the Amazon region. *Journal of Hydrology*, *522*, 674–683. https://doi.org/10.1016/J.JHYDROL.2015.01.021

Lobato, T. C., Hauser-Davis, R. A., Oliveira, T. F., Silveira, A. M., Silva, H. A. N., Tavares, M. R. M., & Saraiva, A. C. F. (2015b). Construction of a novel water quality index and quality indicator for reservoir water quality evaluation: A case study in the Amazon region. *Journal of Hydrology*, *522*, 674–683. https://doi.org/10.1016/j.jhydrol.2015.01.021

Lu, F., Chen, Z., & Liu, W. (2014). A Gis-based system for assessing marine water quality around offshore platforms. *Ocean & Coastal Management*, *102*(PA), 294–306. https://doi.org/10.1016/J.OCECOAMAN.2014.10.003

Lu, G. Y., & Wong, D. W. (2008). An adaptive inverse-distance weighting spatial interpolation technique. *Computers & Geosciences*, *34*(9), 1044–1055. https://doi.org/10.1016/J.CAGEO.2007.07.010

Lumb, A., Sharma, · T C, Bibeault, J.-F., Sharma, T. C., & Bibeault, J.-F. (2011). A Review of Genesis and Evolution of Water Quality Index (WQI) and Some Future Directions. *Water Quality, Exposure and Health 2011 3:1*, *3*(1), 11–24. https://doi.org/10.1007/S12403-011-0040-0

Lumb, A., Sharma, T. C., & Bibeault, J.-F. (2011). A Review of Genesis and Evolution of Water Quality Index (WQI) and Some Future Directions. *Water Quality, Exposure and Health*, *3*(1), 11–24. https://doi.org/10.1007/s12403-011-0040-0

Ma, Z., Li, H., Ye, Z., Wen, J., Hu, Y., & Liu, Y. (2020). Application of modified water quality index (WQI) in the assessment of coastal water quality in main aquaculture areas of Dalian, China. *Marine Pollution Bulletin*, *157*, 111285. https://doi.org/10.1016/J.MARPOLBUL.2020.111285

Mahapatra, S. S., Sahu, M., Patel, · R K, Biranchi, ·, Panda, N., Sahu, M., Patel, R. K., & Panda, B. N. (2012). Prediction of Water Quality Using Principal Component Analysis.

Water Quality, Exposure and Health 2012 4:2, *4*(2), 93–104. https://doi.org/10.1007/S12403-012-0068-9

Mandal, U. K., Warrington, D. N., Bhardwaj, A. K., Bar-Tal, A., Kautsky, L., Minz, D., & Levy, G. J. (2008). Evaluating impact of irrigation water quality on a calcareous clay soil using principal component analysis. *Geoderma*, *144*(1–2), 189–197. https://doi.org/10.1016/j.geoderma.2007.11.014

Maskooni, E. K., Naseri-Rad, M., Berndtsson, R., & Nakagawa, K. (2020). Use of Heavy Metal Content and Modified Water Quality Index to Assess Groundwater Quality in a Semiarid Area. *Water 2020, Vol. 12, Page 1115*, *12*(4), 1115. https://doi.org/10.3390/W12041115

Matthew, C., Lawoko, C. R. O., Korte, C. J., & Smith, D. (1994). Application of canonical discriminant analysis, principal component analysis, and canonical correlation analysis as tools for evaluating differences in pasture botanical composition. *New Zealand Journal of Agricultural Research*, *37*(4), 509–520. https://doi.org/10.1080/00288233.1994.9513090

McGarigal, K., Stafford, S., & Cushman, S. (2000). Multivariate Statistics for Wildlife and Ecology Research. *Multivariate Statistics for Wildlife and Ecology Research*. https://doi.org/10.1007/978-1-4612-1288-1

Ministry of Water Resource. (2001). *National Water Management Plan*.

Moog, O., & Scientific Conference Rivers in the Hindu Kush-Himalaya - Ecology & Environmental Assessment 2008 Kāṭhamāṇḍū; Dhulikhel. (n.d.). *ASSESS-HKH a product of ASSESS-HKH, development of an assessment system to evaluate the ecological status of rivers in the Hindu Kush-Himalayan region ; [proceedings of the Scientific Conference Rivers in the Hindu Kush-Himalaya—Ecology & Environmental Assessment]*.

Mukate, S., Wagh, V., Panaskar, D., Jacobs, J. A., & Sawant, A. (2019). Development of new integrated water quality index (IWQI) model to evaluate the drinking suitability of water. *Ecological Indicators*, *101*, 348–354. https://doi.org/10.1016/J.ECOLIND.2019.01.034

Naik, M. R., Mahanty, B., Sahoo, S. K., Jha, V. N., & Sahoo, N. K. (2022). Assessment of groundwater geochemistry using multivariate water quality index and potential health risk in industrial belt of central Odisha, India. *Environmental Pollution*, *303*, 119161. https://doi.org/10.1016/j.envpol.2022.119161

Nandini, N., Jumbe, A. S., Tandon, S., & N., S. (2007). Application of Remote Sensing, GIS, and GPS Technology on Environment and Earth's Natural Resources Management An Introduction. *Mapana Journal of Sciences*, *6*(1), 96–117. https://doi.org/10.12723/MJS.10.8

Nath, B. K., Chaliha, C., Bhuyan, B., Kalita, E., Baruah, D. C., & Bhagabati, A. K. (2018). GIS mapping-based impact assessment of groundwater contamination by arsenic and other heavy metal contaminants in the Brahmaputra River valley: A water quality assessment study. *Journal of Cleaner Production*, *201*, 1001–1011. https://doi.org/10.1016/J.JCLEPRO.2018.08.084

Nath Roy, B., Roy, H., Rahman, K. S., Mahmud, F., Bhuiyan, M. M. K., Hasan, M., Bhuiyan, A.-A. K., Hasan, M., Mahbub, M. S., Jahedi, R. M., & Islam, M. S. (2024). Principal component analysis incorporated water quality index modeling for Dhaka-based rivers. *City and Environment Interactions*, *23*, 100150. https://doi.org/10.1016/j.cacint.2024.100150

Naubi, I., Zardari, N. H., Shirazi, S. M., Ibrahim, N. F. B., & Baloo, L. (2016). Effectiveness of Water Quality Indexfor Monitoring Malaysian RiverWater Quality. *Polish Journal of Environmental Studies*, *25*(1), 231–239. https://doi.org/10.15244/PJOES/60109

Naushad Alam, M., Fazle Elahi, P., & Md Didar-Ul-Alam, P. (n.d.). *Risk and Water Quality Assessment over view of River Sitalakhya in Bangladesh*.

Navarro, A., Tauler, R., Lacorte, S., & Barceló, D. (2010). Occurrence and transport of pesticides and alkylphenols in water samples along the Ebro River Basin. *Journal of Hydrology*, *383*(1–2), 18–29. https://doi.org/10.1016/J.JHYDROL.2009.06.039

Nazma Sultana, M., Sahadat Hossain, M., & Ara Latifa, G. (2019). WATER QUALITY ASSESSMENT OF BALU RIVER, DHAKA BANGLADESH. *Water Conservation & Management (WCM)*. https://doi.org/10.26480/wcm.02.2019.08.10

Nong, X., Shao, D., Zhong, H., & Liang, J. (2020). Evaluation of water quality in the South-to-North Water Diversion Project of China using the water quality index (WQI) method. *Water Research*, *178*, 115781. https://doi.org/10.1016/J.WATRES.2020.115781

Noori, R., Berndtsson, R., Hosseinzadeh, M., Adamowski, J. F., & Abyaneh, M. R. (2019). A critical review on the application of the National Sanitation Foundation Water Quality Index. *Environmental Pollution*, *244*, 575–587. https://doi.org/10.1016/J.ENVPOL.2018.10.076

Oketola, A. A., Adekolurejo, S. M., Osibanjo, O., Oketola, A. A., Adekolurejo, S. M., & Osibanjo, O. (2013). Water Quality Assessment of River Ogun Using Multivariate Statistical Techniques. *Journal of Environmental Protection*, *4*(5), 466–479. https://doi.org/10.4236/JEP.2013.45055

Olsen, R. L., Chappell, R. W., & Loftis, J. C. (2012). Water quality sample collection, data treatment and results presentation for principal components analysis – literature review and Illinois River watershed case study. *Water Research*, *46*(9), 3110–3122. https://doi.org/10.1016/J.WATRES.2012.03.028

Omer, N. H., & Omer, N. H. (2019). Water Quality Parameters. *Water Quality - Science, Assessments and Policy*. https://doi.org/10.5772/INTECHOPEN.89657

Oseke, F. I., Anornu, G. K., Adjei, K. A., & Eduvie, M. O. (2021). Assessment of water quality using GIS techniques and water quality index in reservoirs affected by water diversion. *Water-Energy Nexus*, *4*, 25–34. https://doi.org/10.1016/J.WEN.2020.12.002

Ouyang, Y. (2005). Evaluation of river water quality monitoring stations by principal component analysis. *Water Research*, *39*(12), 2621–2635. https://doi.org/10.1016/J.WATRES.2005.04.024

Ouyang, Y., Nkedi-Kizza, P., Wu, Q. T., Shinde, D., & Huang, C. H. (2006). Assessment of seasonal variations in surface water quality. *Water Research*, *40*(20), 3800–3810. https://doi.org/10.1016/J.WATRES.2006.08.030

Parmar, K. S., & Bhardwaj, R. (2014). Water quality management using statistical analysis and time-series prediction model. *Applied Water Science*, *4*(4), 425–434. https://doi.org/10.1007/S13201-014-0159-9/FIGURES/3

Peche, R., & Rodríguez, E. (2012). Development of environmental quality indexes based on fuzzy logic. A case study. *Ecological Indicators*, *23*, 555–565. https://doi.org/10.1016/J.ECOLIND.2012.04.029

Pesce, S. F., & Wunderlin, D. A. (2000). Use of water quality indices to verify the impact of Córdoba City (Argentina) on Suquía River. *Water Research*, *34*(11), 2915–2926. https://doi.org/10.1016/S0043-1354(00)00036-1

Pinto, L. C., de Mello, C. R., & Ávila, L. F. (2013). Water quality indicators in the Mantiqueira Range region, Minas Gerais state. *CERNE*, *19*(4), 687–692. https://doi.org/10.1590/S0104-77602013000400020

Platikanov, S., Rodriguez-Mozaz, S., Huerta, B., Barceló, D., Cros, J., Batle, M., Poch, G., & Tauler, R. (2014). Chemometrics quality assessment of wastewater treatment plant effluents using physicochemical parameters and UV absorption measurements. *Journal of Environmental Management*, *140*, 33–44. https://doi.org/10.1016/J.JENVMAN.2014.03.006

Principal Component Analysis for Special Types of Data. (2002). *Principal Component Analysis*, 338–372. https://doi.org/10.1007/0-387-22440-8_13

Rahman, A., Islam, M., Hossain, M., & Ahsan, M. A. (2012). Study of the seasonal variations in Turag river water quality parameters. *African Journal of Pure and Applied Chemistry*, *6*(10). https://doi.org/10.5897/AJPAC12.023

Rahman, M. A., & Al Bakri, D. (2010). A Study on Selected Water Quality Parameters along the River Buriganga, Bangladesh. *Official Peer Reviewed Journal of Babol Noshirvani University of Technology BUT*, *1*(2), 81–92.

Rahman, S., & Hossain, F. (2008). Spatial assessment of water quality in Peripheral Rivers of Dhaka City for optimal relocation of water intake point. *Water Resources Management*, *22*(3), 377–391. https://doi.org/10.1007/S11269-007-9167-Y/METRICS

Rai, R. K., Upadhyay, A., Ojha, C. S. P., & Singh, V. P. (2012). *The Yamuna River Basin* (Vol. 66). Springer Netherlands. https://doi.org/10.1007/978-94-007-2001-5

Ringnér, M. (2008). What is principal component analysis? *Nature Biotechnology 2008 26:3*, *26*(3), 303–304. https://doi.org/10.1038/nbt0308-303

Rocha, F. C., Andrade, E. M., & Lopes, F. B. (2015). Water quality index calculated from biological, physical and chemical attributes. *Environmental Monitoring and Assessment*, *187*(1), 1–15. https://doi.org/10.1007/S10661-014-4163-1/TABLES/9

Roy, R. (2018). An Introduction to water quality analysis. *ESSENCE – International Journal for Environmental Rehabilitation and Conservation*, 94–100. https://doi.org/10.31786/09756272.18.9.2.214

Rusydi, A. F. (2018). Correlation between conductivity and total dissolved solid in various type of water: A review. *IOP Conference Series: Earth and Environmental Science*, *118*, 012019. https://doi.org/10.1088/1755-1315/118/1/012019

Sabinaya, S., Mahanty, B., Rout, P. R., Raut, S., Sahoo, S. K., Jha, V., & Sahoo, N. K. (2024). Multi-model exploration of groundwater quality and potential health risk assessment in Jajpur district, Eastern India. *Environmental Geochemistry and Health*, *46*(2), 57. https://doi.org/10.1007/s10653-024-01855-1

Said, A., Stevens, D. K., & Sehlke, G. (2004). An innovative index for evaluating water quality in streams. *Environmental Management*, *34*(3), 406–414. https://doi.org/10.1007/S00267-004-0210-Y/FIGURES/3

Salman, M. A., Ahmed, S., Peas, M. H., & Khan, N. (2018). Water quality assessment of the Buriganga river, Dhaka, Bangladesh. *International Journal of Emerging Technology and Advanced Engineering*, *8*(6), 17–23.

Sánchez, E., Colmenarejo, M. F., Vicente, J., Rubio, A., García, M. G., Travieso, L., & Borja, R. (2007). Use of the water quality index and dissolved oxygen deficit as simple indicators of watersheds pollution. *Ecological Indicators*, *7*(2), 315–328. https://doi.org/10.1016/J.ECOLIND.2006.02.005

Sarkar, C., & Abbasi, S. A. (2006a). Qualidex – A New Software for Generating Water Quality Indice. *Environmental Monitoring and Assessment*, *119*(1–3), 201–231. https://doi.org/10.1007/s10661-005-9023-6

Sarkar, C., & Abbasi, S. A. (2006b). Qualidex—A new software for generating water quality indice. *Environmental Monitoring and Assessment*, *119*(1–3), 201–231. https://doi.org/10.1007/S10661-005-9023-6/METRICS

Sayed Ahammed, S., Tasfina, S., Ayaz Rabbani, K., & Adbul Khaleque, M. (2016). An Investigation Into The Water Quality Of Buriganga-A River Running Through Dhaka. *INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH*, *5*(03).

Shah, K. A., & Joshi, G. S. (2017). Evaluation of water quality index for River Sabarmati, Gujarat, India. *Applied Water Science*, *7*(3), 1349–1358. https://doi.org/10.1007/S13201-015-0318-7/TABLES/5

Sharma, A., Singh, A. K., & Kumar, K. (2012). Environmental geochemistry and quality assessment of surface and subsurface water of Mahi River basin, western India. *Environmental Earth Sciences*, *65*(4), 1231–1250. https://doi.org/10.1007/S12665-011-1371-7/FIGURES/11

Shibata, T., Solo-Gabriele, H. M., Fleming, L. E., & Elmir, S. (2004). Monitoring marine recreational water quality using multiple microbial indicators in an urban tropical environment. *Water Research*, *38*(13), 3119–3131. https://doi.org/10.1016/J.WATRES.2004.04.044

Shrestha, S., & Kazama, F. (2007). Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan. *Environmental Modelling & Software*, *22*(4), 464–475. https://doi.org/10.1016/J.ENVSOFT.2006.02.001

Singh, A. K., Mondal, G. C., Kumar, S., Singh, T. B., Tewary, B. K., & Sinha, A. (2008). Major ion chemistry, weathering processes and water quality assessment in upper catchment of Damodar River basin, India. *Environmental Geology*, *54*(4), 745–758. https://doi.org/10.1007/S00254-007-0860-1/FIGURES/13

Singh, K. P., Malik, A., Mohan, D., & Sinha, S. (2004). Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India)—A case study. *Water Research*, *38*(18), 3980–3992. https://doi.org/10.1016/J.WATRES.2004.06.011

Singh, S. K., Singh, P., & Gautam, S. K. (2016). Appraisal of urban lake water quality through numerical index, multivariate statistics and earth observation data sets. *International Journal of Environmental Science and Technology*, *13*(2), 445–456. https://doi.org/10.1007/S13762-015-0850-X/TABLES/6

Stoner, J. D. (1978). *Water-quality Indices for Specific Water Uses*.

Stoner, J. D., Andrus, C. D., & Menard, H. W. (n.d.). *Water-Quality Indices for Specific Water Uses United States Department of the Interior Geological Survey*.

Summers, K., & Summers, K. (2020). Water Quality—Science, Assessments and Policy. *Water Quality - Science, Assessments and Policy*. https://doi.org/10.5772/INTECHOPEN.77531

Sun, W., Xia, C., Xu, M., Guo, J., & Sun, G. (2016). Application of modified water quality indices as indicators to assess the spatial and temporal trends of water quality in the Dongjiang River. *Ecological Indicators*, *66*, 306–312. https://doi.org/10.1016/J.ECOLIND.2016.01.054

Sutadian, A. D., Muttil, N., Yilmaz, A. G., & Perera, B. J. C. (2017). Using the Analytic Hierarchy Process to identify parameter weights for developing a water quality index. *Ecological Indicators*, *75*, 220–233. https://doi.org/10.1016/J.ECOLIND.2016.12.043

Sutadian, A. D., Muttil, N., Yilmaz, A. G., & Perera, B. J. C. (2018a). Development of a water quality index for rivers in West Java Province, Indonesia. *Ecological Indicators*, *85*, 966–982. https://doi.org/10.1016/J.ECOLIND.2017.11.049

Sutadian, A. D., Muttil, N., Yilmaz, A. G., & Perera, B. J. C. (2018b). Development of a water quality index for rivers in West Java Province, Indonesia. *Ecological Indicators*, *85*, 966–982. https://doi.org/10.1016/j.ecolind.2017.11.049

Swamee, P. K., & Tyagi, A. (2000). Describing Water Quality with Aggregate Index. *Journal of Environmental Engineering*, *126*(5), 451–455. https://doi.org/10.1061/(ASCE)0733-9372(2000)126:5(451)

Swamee, P. K., & Tyagi, A. (2007). Improved Method for Aggregation of Water Quality Subindices. *Journal of Environmental Engineering*, *133*(2), 220–225. https://doi.org/10.1061/(ASCE)0733-9372(2007)133:2(220)/ASSET/D70772CB-9778-4BC0-AEE9-FED8EAA111BA/ASSETS/IMAGES/LARGE/6.JPG

Tan, G., Yan, J., Gao, C., & Yang, S. (2012). Prediction of water quality time series data based on least squares support vector machine. *Procedia Engineering*, *31*, 1194–1199. https://doi.org/10.1016/J.PROENG.2012.01.1162

Tharmar, E., Abraham, M., Prakash, R., Sundaram, A., Flores, E. S., Canales, C., & Alam, M. A. (2022). Hydrogeochemistry and Water Quality Assessment in the Thamirabarani River Stretch by Applying GIS and PCA Techniques. *Sustainability 2022, Vol. 14, Page 16368*, *14*(24), 16368. https://doi.org/10.3390/SU142416368

Thi Minh Hanh, P., Sthiannopkao, S., The Ba, D., & Kim, K.-W. (2011). Development of Water Quality Indexes to Identify Pollutants in Vietnam's Surface Water. *Journal of Environmental Engineering*, *137*(4), 273–283. https://doi.org/10.1061/(ASCE)EE.1943-7870.0000314

Tirkey, P., Bhattacharya, T., Chakraborty, S., Poonam, T., Tanushree, B., & Sukalyan, C. (2013). Water quality indices-important tools for water quality assessment: A review. *International Journal of Advances in Chemistry (IJAC)*, *1*(1).

Tomas, D., Čurlin, M., & Marić, A. S. (2017). Assessing the surface water status in Pannonian ecoregion by the water quality index model. *Ecological Indicators*, *79*, 182–190. https://doi.org/10.1016/J.ECOLIND.2017.04.033

Toming, K., Kutser, T., Laas, A., Sepp, M., Paavel, B., & Nõges, T. (2016). First Experiences in Mapping Lake Water Quality Parameters with Sentinel-2 MSI Imagery. *Remote Sensing 2016, Vol. 8, Page 640*, *8*(8), 640. https://doi.org/10.3390/RS8080640

Tripathi, M., & Singal, S. K. (2019). Use of Principal Component Analysis for parameter selection for development of a novel Water Quality Index: A case study of river Ganga

India. *Ecological Indicators*, *96*, 430–436. https://doi.org/10.1016/j.ecolind.2018.09.025

Uddin, G., Nash, S., & Olbert, A. I. (2022). *Optimization of parameters in a water quality index model using principal component analysis*. 5739–5744. https://doi.org/10.3850/IAHR-39WC2521711920221326

Uddin, M. G., Nash, S., & Olbert, A. I. (2021). A review of water quality index models and their use for assessing surface water quality. *Ecological Indicators*, *122*, 107218. https://doi.org/10.1016/J.ECOLIND.2020.107218

Uddin, Md. G., Moniruzzaman, Md., & Khan, M. (2017). Evaluation of Groundwater Quality Using CCME Water Quality Index in the Rooppur Nuclear Power Plant Area, Ishwardi, Pabna, Bangladesh. *American Journal of Environmental Protection, Vol. 5, 2017, Pages 33-43*, *5*(2), 33–43. https://doi.org/10.12691/ENV-5-2-2

Ustaoğlu, F., Tepe, Y., & Taş, B. (2020). Assessment of stream quality and health risk in a subtropical Turkey river system: A combined approach using statistical analysis and water quality index. *Ecological Indicators*, *113*, 105815. https://doi.org/10.1016/J.ECOLIND.2019.105815

Uthicke, S., Vogel, N., Doyle, J., Schmidt, C., & Humphrey, C. (2012). Interactive effects of climate change and eutrophication on the dinoflagellate-bearing benthic foraminifer Marginopora vertebralis. *Coral Reefs*, *31*(2), 401–414. https://doi.org/10.1007/S00338-011-0851-2/FIGURES/5

Wang, L., Mei, Y., Yu, K., Li, Y., Meng, X., & Hu, F. (2019). Anthropogenic Effects on Hydrogeochemical Characterization of the Shallow Groundwater in an Arid Irrigated Plain in Northwestern China. *Water 2019, Vol. 11, Page 2247*, *11*(11), 2247. https://doi.org/10.3390/W11112247

Whitehead, P., Bussi, G., Hossain, M. A., Dolk, M., Das, P., Comber, S., Peters, R., Charles, K. J., Hope, R., & Hossain, S. (2018). Restoring water quality in the polluted Turag-Tongi-Balu river system, Dhaka: Modelling nutrient and total coliform intervention strategies. *Science of The Total Environment*, *631–632*, 223–232. https://doi.org/10.1016/J.SCITOTENV.2018.03.038

WHO. (2024). *Guidelines for drinking-water quality: Small water supplies*.

Wu, C. Y., Mossa, J., Mao, L., & Almulla, M. (2019). Comparison of different spatial interpolation methods for historical hydrographic data of the lowermost Mississippi River. *Annals of GIS*, *25*(2), 133–151. https://doi.org/10.1080/19475683.2019.1588781

Wunderlin, D. A., María Del Pilar, D., María Valeria, A., Fabiana, P. S., Cecilia, H. A., & María De Los Ángeles, B. (2001). Pattern Recognition Techniques for the Evaluation of Spatial and Temporal Variations in Water Quality. A Case Study: Suquía River Basin (Córdoba–Argentina). *Water Research*, *35*(12), 2881–2894. https://doi.org/10.1016/S0043-1354(00)00592-3

Xia, R., & Chen, Z. (2015). Integrated Water-Quality Assessment of the Huai River Basin in China. *Journal of Hydrologic Engineering*, *20*(2), 05014018. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001030/ASSET/2D06B252-D463-4117-8088-D8E15E4CA10D/ASSETS/IMAGES/LARGE/FIGURE3.JPG

Yang, Q., Hou, Z., Wang, Y., Zhao, Y., & Delgado, J. (2015). A comparative study of shallow groundwater level simulation with WA–ANN and ITS model in a coastal island of south China. *Arabian Journal of Geosciences*, *8*(9), 6583–6593. https://doi.org/10.1007/S12517-014-1706-2/TABLES/1

Yang, Q., Li, Z., Ma, H., Wang, L., & Martín, J. D. (2016). Identification of the hydrogeochemical processes and assessment of groundwater quality using classic integrated geochemical methods in the Southeastern part of Ordos basin, China. *Environmental Pollution*, *218*, 879–888. https://doi.org/10.1016/J.ENVPOL.2016.08.017

Yang, Q., Zhang, J., Wang, Y., Fang, Y., & Martín, J. D. (2015). Multivariate Statistical Analysis of HydrochemicalData for Shallow Ground Water Quality FactorIdentification in a Coastal Aquifer. *Polish Journal of Environmental Studies*, *24*(2), 769–776. https://doi.org/10.15244/PJOES/30263

Zou, H., Hastie, T., & Tibshirani, R. (2006). Sparse Principal Component Analysis. *Journal of Computational and Graphical Statistics*, *15*(2), 265–286. https://doi.org/10.1198/106186006X113430

Chapter 12: Appendices

Appendix A: Agreement Signing

The agreement between Dept. of Chemical Engineering, BUET and the Water Resources Planning Organization (WARPO) to conduct the collaborative research project on "Establishment of Water Quality Index (WQI) through Principal Component Analysis for the Dhaka-based Rivers" was signed on 06 June 2023 (Figure 12.1). The Director General of WARPO, along with senior scientific officers and other officials, were present at this ceremony. On behalf of the Research team of the Dept. of ChE, the Principal Investigator, Professor Dr. Md. Shahinoor Islam was also present at this event. The ceremony was held in WARPO Bhaban, Green Road, Dhaka.

Figure 12.1: Contract signing between WARPO and Dept. of ChE, BUET

Appendix B: Field Visit

The first sampling was conducted on Friday, 21st July, 2023. Members from the BUET team, including the principal investigator Dr. Md Shahinoor Islam and members from the WARPO team were present in the sample collection. The sample was collected from the BG4 station. A picture of the members present at the first sample collection is shown in Figure 12.2.

Figure 12.2: Members present in sample collection at Buriganga river

Another field visit was conducted on Friday, 04th August 2023, at the TR1 station in Turag river, near Fulpukuria Thread & Accessories Ltd. Images taken on the second sample collection are shown in Figure 12.3. A field visit and sample collection were conducted at the Balu river near Balu bridge, 300 feet, Dhaka. Figure 12.4 contains the images from the field visit at Balu Bridge on Wednesday, 06th September 2023. On 28th September another field visit was conducted in Ghorashal Gudara ghat at the SL1 station in Shitalakhya river. Figure 12.5 contains the images from the field visit at Ghorashal point.

Figure 12.3: Members present in sample collection at Turag river

Figure 12.4: Members present in sample collection at at Balu river

Figure 12.5: Members present in sample collection at at Shitalakhya river

Appendix C: Technical Committee Meeting: December 2023

On December 14th, 2023 a technical committee meeting (TCM) was held at the WARPO seminar room. Md. Rezaul Maksud Jahedi, Director General, WARPO was the chairperson of the meeting. Figure 12.6 was a captured moment of the 1st TCM.

Figure 12.6: Technical committee meeting: December 2023 held at WARPO

Appendix D: Office and Lab Setup

After contract signing the Department of ChE, BUET has provided an office room in the post graduate hub in the department as shown in Figure 12.7. Sufficient accommodation along with logistic facilities have also been provided.

Figure 12.7: Office setup at Department of Chemical Engineering, BUET

Also, after signing the contract, four lab equipment were bought for water sample analysis which are currently placed in the Air Pollution Lab in the Department of ChE, BUET and will be added to WARPO lab facilities after the construction is finished. The following are the lab equipment bought for this project.

Portable Multiparameter Water Quality Meter: Hanna HI98194

This three-probe multimeter shown in Figure 12.8 can measure on-site temperature, pressure, pH, electrical conductivity (EC), total dissolved solids (TDS), dissolved oxygen (DO), salinity, oxygen reduction potential (ORP). Being portable, this device is able to provide precise and accurate measurements on-site. After immersing the cylindrical device with three probes in the sample water and holding it still for some time, all the parameters are available on the meter screen.

Figure 12.8: Portable Multiparameter Water Quality Meter

UV-VIS Spectrophotometer: SCAN Spectro::lyser V3

The equipment shown in Figure 12.9 is a state-of-the-art UV-Vis spectrometer that can measure on-site Biological Oxygen Demand (BOD5), Chemical Oxygen Demand (COD), Total Suspended Solids (TSS), Dissolved Organic Carbon (DOC), Total Organic Carbon (TOC). BOD5 determination with conventional technology takes 5 days, requires extensive lab work, and TOC determination is quite difficult. This state-of-the-art Spectro::lyser V3 can be a new addition to water research in Bangladesh, which is time and cost efficient at the same time. Various parameters can be determined with the help of spectrophotometry through this spectrometer with titanium probe. With the help of on-site battery, it is possible to easily view the readings on the mobile screen by connecting to the smart phone via WiFi hotspot while using this probe.

Figure 12.9: UV-VIS Spectrophotometer

Portable Spectrophotometer: HACH DR-1900

The portable spectrophotometer shown in Figure 12.10 can measure chloride, nitrate-nitrogen, phosphate, ammonia-nitrogen and other water quality parameters using specific kits. It is suitable for both on-site and lab use. Phosphate, chloride, nitrate and ammonium-nitrogen can be determined in a very short time by entering 10 ml of blank sample and 10 ml of reagent mixed sample into two sample cells by selecting the specified program in the spectrophotometer.

Figure 12.10: Portable Spectrophotometer

Most Probable Number (MPN) Analyzer: Microbium MPN Analyser Basic – L

With the help of the modern incubator and kit shown in Figure 12.11, E. Coli and Total Coliform in water can be easily detected within 24 hours and their Most Probable Number (MPN) is obtained. Prior to the invention of this sophisticated instrument, E. Coli and Total Coliform determinations were analog, time consuming (critical counting), and labor intensive. This instrument is capable of adding a new dimension to water research in Bangladesh. 24 holes of the MPN kit are filled with 100 ml of reagent-mixed sample and kept in the incubator at 36°C for one day to show different color combinations. From there, by inputting the color combination to the analyzing software, E.Coli. and the amount of Total Coliform is obtained.

Figure 12.11: Most Probable Number (MPN) Analyzer

Appendix E: Technical Workshop: May 08-09, 2024

A two-days long technical workshop was held at the WARPO seminar room from May 08, 2024 to May 09, 2024. This workshop was conducted by Dr. Md. Shahinoor Islam, Professor and Head, DChE, BUET and Hridoy Roy, Assistant Professor, DChE, BUET. This workshop was designed in four modules and each day two modules were conducted. On the first day, modules, Water Quality parameters and Water Quality Assessment and on the second day Water Quality Index and Case Study – Integrating PCA into WQI Model were conducted. Md. Rezaul Maksud Jahedi, Director General, WARPO inaugurated this workshop and was present on both days. After finishing the workshop, certificates were provided for each participant. Figure 12.12 shows some snapshots from the technical workshop.

Figure 12.12: Technical workshop held on WARPO: May 08-09, 2024

Appendix F: Technical Workshop: May 06, 2025

A one-day technical workshop was held at the WARPO Water Analysis Laboratory on May 6, 2025. The workshop was conducted by Hridoy Roy, Assistant Professor in the Department of Chemical Engineering, BUET. It was organized in two modules. The first module focused on Principal Component Analysis (PCA), covering both the theoretical foundation and practical implementation using R. The second module addressed the development of the Water Quality Index (WQI), including all four main parts of WQI formulation.

Figure 12.13: Technical workshop held on WARPO: May 06, 2025

Appendix G: Technical Committee Meeting: May 2025

A Technical Committee Meeting (TCM) was held at the WARPO office on 15 May 2025 at 10:30 AM. The meeting was chaired by Mohammad Lutfur Rahman, Director General of WARPO. In the meeting, Md. Shahinoor Islam from the Department of Chemical Engineering, BUET, presented the draft final report. The purpose of the meeting was to review the technical progress and gather feedback from the committee members.

Figure 12.14: Technical committee meeting: May 2025 held at WARPO

Appendix H: Project Activity Timeline

Table 12.1: Activity Schedule starting from June, 2023

Sl	Description of Activities	Months																							
		M 1	M 2	M 3	M 4	M 5	M 6	M 7	M 8	M 9	M 10	M 11	M 12	M 13	M 14	M 15	M 16	M 17	M 18	M 19	M 20	M 21	M 22	M 23	M 24
1.	Mobilization	■	■																						
2.	Site Selection		■	■	■					■	■														
3.	Site Visits and Secondary Data Collection and Analysis		■	■	■	■	■	■	■																
4.	Sample Collection				■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■				
5.	Lab Analysis						■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■			
6.	Principal Component Analysis								■	■	■	■	■	■	■	■	■	■	■	■	■	■	■		
7.	Determination of WQI										■	■	■	■	■	■	■	■	■	■	■	■	■		
8.	Comparison of WQIs																					■	■		
9.	WQI Establishment																						■	■	

Sl	Description of Activities	Months																							
		M 1	M 2	M 3	M 4	M 5	M 6	M 7	M 8	M 9	M 10	M 11	M 12	M 13	M 14	M 15	M 16	M 17	M 18	M 19	M 20	M 21	M 22	M 23	M 24
10.	Inception Report																								
11.	Technical Committee Meeting																								
12.	Interim Report																								
13.	Draft Final Report																								
14.	Final Workshop																								
15.	Final Report																								

Continuous

Intermittent

Appendix I: Research Proposal

The following research proposal was presented to WARPO authority.

Collaborative Research Project Brief

Establishment of Water Quality Index (WQI) through Principal Component Analysis for the Dhaka-based Rivers

1

BACKGROUND:

Rapid urbanization has been a boon for industrial growth in Bangladesh, leading Dhaka, the capital and center of trade and business, to be one of the least livable places in the world. Dhaka, the home of more than 20 million people and one of the most densely populated megacities in the world, remains at the bottom in terms of livable cities. It has grown along the rivers Buriganga, Turag, Balu and Shitalakshya for the past 400 years (Figure 1). The unplanned development and much of its urbanization are linked to its industrial development, which over time has resulted in decreased water quality of the rivers.

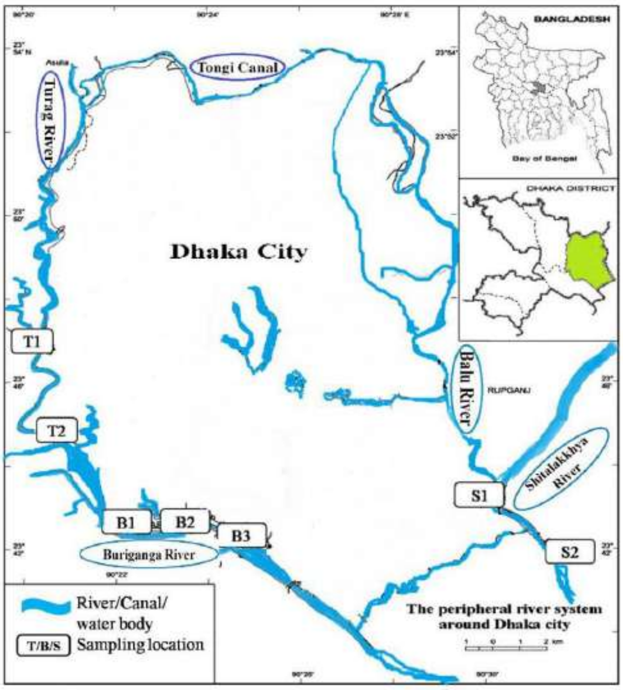


Figure 1. Buriganga, Turag, Balu and Shitalakshya rivers on the map of Dhaka

Inadequate sewage and inefficient industrial waste management contributed to river pollution, resulting in water quality in the river to parameters far below the critical limit. Moreover, untreated industrial waste and household sewage are discharged into the river system (after ineffective or without treatment), leading to the extinction of aquatic life and the failure of ecosystems [1]. The mismanagement and lack of awareness of the people of Dhaka have turned Buriganga, Turag, Balu and Shitalakshya into a polluted storage basin (Figure 2).

2

Figure 2: The pollution scenario of Buriganga, Turag, Balu and Shitalakshya rivers

There are some studies available on the Dhaka-based river water quality assessment. Mir Mostafa Kamal et. al. (1999) reported the alarmingly low level of DO in the Buriganga river back in the 2000s [2]. Many other researchers studied water quality parameters e.g., dissolved oxygen (DO), biochemical oxygen demand (BOD), chemical oxygen demand (COD), pH, turbidity, conductivity, total dissolved solids (TDS), nitrate and phosphate in Buriganga river and those studies found the DO, BOD, COD, TDS, turbidity, nitrate and phosphate are at an alarming level and a discussion on the possible sources of the pollution are presented in some of the papers [3]. Paul Whitehead et. al. (2018) and his team did a baseline survey of water chemistry and total coliforms of the Turag-Tongi-Balu river system and showed DO close to zero in the dry season, high organic loading together with extreme levels of Ammonium-N and total coliform in the water [4]. However, most of the studies are scattered, non-comprehensive, dealt with several parameters and limited sites of the river. These studies emphasize more on insights, causes and consequences of river water pollution. Thus, these studies could not reach the policymakers. In order to introduce a long-term and effective river water management, an oriented and structured research is necessary to understand the actual condition of the study area (baseline). In order to recognize the Dhaka-based river water quality, a good source of data on the water environment is indispensable. Although the regular measurements need much work, because of spatial and temporal variation of water environment quality, monitoring by regular measurements, which would provide a representative and reliable estimation of surface water quality, is necessary. The long-term monitoring of many profiles with different reach would generate a large and complex database, which needs a good approach to interpretation [5].

To mitigate and trace the pollution source, the correct interpretation of the collected data is crucial. Moreover, random site selection of study areas is also an issue in water quality assessment. The geographical information system (GIS) is very helpful to evaluate the spatial distribution of water quality parameters over the study area [6]. The collected data can be analyzed by the application of different multivariate statistical techniques, such as Principal Component Analysis (PCA), Cluster Analysis (CA), and multiple linear regression. These techniques help in the interpretation of complex data matrices to better understand the water quality and ecological status of the studied systems and allow the identification of possible factors that influence water environment systems and offer a valuable tool for the reliable management of water resources. Assessment of river water quality or optimization of the monitoring procedure of river water is linked to clustering of sampling locations, river water quality parameters, identification of possible sources of pollution, or modeling the contribution of the identified sources to the formation of the total concentration of the monitored chemical tracers.

Most of the time, the research on water quality assessment fails due to the thousands of output information because water quality regulations and monitoring programs generate a vast number of multidimensional water quality parameter sets, pollution characteristics, and numerical data about various water sources that are understandable only to the scientists [7]. This information, however, should be beneficial to water sector managers in making decisions who want to update the condition of their water sources. As a result, a technique known as the Water Quality Index (WQI) should be developed to solve this issue. The Water Quality Index is a quantitative representation that is used to determine the ecological health of a body of water. The purpose of the Water Quality Index is to categorize waters according to their physical, chemical, and biological attributes, thereby establishing their potential uses and controlling their allocation decisions [16]. By providing a single dimensionless value, the WQI helps to reduce the multivariate nature of the data that describes the quality of water [8].

Thus, proper site selection, correct analysis of the collected data, and indicative results all should be a part of a river water quality assessment project. Otherwise, the research would be limited to the researcher only. In this research, a comprehensive research would be carried out to understand four main river water characteristics using ArcGIS, numerical Principal Component Analysis (PCA) and WQI analysis.

HYPOTHESIS:

The Hypothesis of this research is:

"A comprehensive Water Quality Index (WQI) framework can be developed for the rivers around Dhaka city based on the critical physical, chemical and biological parameters to accurately assess the overall water quality of the rivers, and can serve as a useful tool for policymakers and relevant stakeholders to make informed decisions about water quality management."

OBJECTIVE:

The main objective of the research is to establish a comprehensive method to determine Water Quality Index (WQI) for the Dhaka-based rivers. On a pilot basis, the research would concentrate on the selected reaches of the Buriganga, Turag, Balu and Shitalakshya rivers.

The specific objectives are,

- To understand the baseline status and condition of the water quality of the Buriganga, Turag, Balu and Shitalakshya rivers
- To identify the most critical water quality parameters for the selected rivers using principal component analysis (PCA)
- To establish a method/approach, consistent with internationally adopted approach, to calculate water quality index (WQI) for the selected rivers
- To provide a comprehensive guideline and recommendation for future rehabilitation initiatives for these rivers.

OVERALL GOAL:

The overall goal of the research is to achieve a comprehensive understanding of the water quality status of the selected study areas and to support river management through proper planning and design by developing a predictive method to suggest efficient approaches for the improvement of the situation. The outcome of the research would be an asset for WARPO in predicting the future aspects and taking necessary steps for the rehabilitation of the Buriganga, Turag, Balu and Shitalakshya rivers. The specific goals are, i) to understand the water quality of the Buriganga, Turag, Balu and Shitalakshya rivers, which would provide a comprehensive guideline for rehabilitation of these rivers, ii) to determine WQI for Buriganga, Turag, Balu and Shitalakshya rivers.

SCOPE OF WORK:

- Collect secondary water quality data of 4 rivers from National Water Resources Database (NWRD), DoE and other secondary sources
- Select potential monitoring sites for sample collection for each river through field survey
- Collect water samples from these monitoring sites following standard protocols, **covering pre-monsoon, monsoon and dry seasons.**
- Perform laboratory tests for important water quality parameters (pH, EC, TDS, DO, BOD, COD, TC etc.)
- Analyze the parameters by Principal Component Analysis (PCA) to obtain the critical water quality parameters
- Derive an appropriate method to calculate water quality index (WQI) for the selected rivers following internationally adopted WQI approach.

METHODOLOGY:

Data Collection:

The methodology starts with a preliminary analysis of the bi-monthly surface water quality variables for the evaluation of temporal and spatial variations and the interpretation of the concentration of pollutants. This analysis helps to identify the most polluted segments of the four rivers and their adjacent streams, as well as the quality variables with the best and worst performances. This is a valuable step towards designing the more complex and rigorous procedure that we describe in the succeeding paragraphs. Initially, we would select **6 numbers** of monitoring sites for each river based on the literature review of previous reports of WARPO, DoE and IWM, field survey and analysis through the Arc-GIS software. The sample collection frequency would be 2 months during **pre-monsoon, monsoon season and dry season as well as during cultural and religious festivals**. Table 1 shows the different parameters that would be considered for this study.

6

Table 1: Different considerable parameters and analyzing instruments/methods

Sl. No.	Parameter	Unit	Instrument/Method
1	Temperature (°C)	°C	Thermometer
2	pH	-	pH meter
3	Electrical conductivity	microS/cm	Glass Electrode
4	TS	mg/L	Standard Method
5	TDS	mg/L	Glass Electrode
6	DO	mg/L	DO meter
7	BOD	mg/L	Standard Method
8	TSS	mg/L	Standard Method
9	COD	mg/L	Standard Method
10	Chloride	mg/L	Standard Method
11	Total Coliform (TC)	mg/L	Membrane filtration
12	Nitrate (NO_3^-)	mg/L	Standard Method
13	Phosphate (PO_4^{3-})	mg/L	Standard Method
14	Ammonia (NH_3)	mg/L	Standard Method

Analysis of Water Quality Parameters:

Standard testing methods including American Public Health Association (APHA), American Society for Testing and Materials (ASTM), US- Environmental Protection Agency (US-EPA), etc. would be applied for analyzing the samples. The following standard protocols would be used for the sample analysis:

1. **Temperature (°C)**: Thermometric analysis using APHA-2550 protocol
2. **pH:** pH meter-based analysis using ASTM-D-1293 protocol
3. **Electrical conductivity:** Glass electrode-based method
4. **TS:** Gravimetric method APHA 2540 B
5. **TDS:** Gravimetric method APHA 2540 C
6. **DO:** Glass electrode-based method
7. **BOD:** APHA-5210 B protocol
8. **TSS:** Gravimetric method APHA 2540 D protocol
9. **COD:** APHA-5220 D protocol
10. **Chloride:** ASTM-E-201 colorimetric method
11. **Total Coliform (TC):** US-EPA membrane filtration method
12. **Nitrate (NO_3^-):** APHA-4500- NO_3^- B protocol
13. **Phosphate (PO_4^{3-}):** Spectrophotometric method
14. **Ammonia (NH_3):** APHA-4500- NH_3 B protocol

7

Spatial Interpolation Using ArcGIS:

GIS is an indispensable tool for natural resource management, particularly in the areas of land use planning, animal habitat analysis, and natural hazard assessment to name a few of its many applications [9]. Although statistical surfaces, in the sense that they are defined by cartographers, do not exist in the same way that land does, it is possible to conceptualize them in the same way. For the preparation of statistical surfaces, the spatial interpolation process in GIS is commonly utilized [10]. Given the impossibility of collecting field data at every location in the study area, spatial interpolation methods would be used to extrapolate information from sampled locations to locations where it could not be collected. Aiming toward the creation of maps depicting the spatial distribution of water quality parameters, ArcInfo 10 GIS software would be used to correlate the water quality to the sampling locations [11]. There are various spatial interpolation methods available. The inverse distance weighting (IDW) method would be used in this case [12]. By using the spatial interpolation method, maps depicting the spatial distribution of the water quality parameters with specified sites would be created. Identifying variations in the accumulations of various parameters in the surface water of the research area was performed.

Principal Component Analysis:

In this research, we would adopt a statistical methodology to monitor the water quality of Dhaka-based four rivers (Buriganga, Turag, Balu and Shitalakshya). The developed method can be easily applied to other river basins suffering from similar environmental problems. The developed method can rigorously reflect the trends and patterns in water quality and are easy to apply and interpret.

There are several multivariate analysis available to analyze water quality parameter data including Principal Component Analysis (PCA), Cluster Analysis (CA), and multiple linear regression etc. [13]. These techniques help in the interpretation of complex data matrices to better understand the water quality and ecological status of the studied systems and allow the identification of possible factors that influence water environment systems and offer a valuable tool for the reliable management of water resources [14]. Among these techniques, PCA has been received interest in the water quality analysis. PCA can be used to identify the critical parameters that severely affect the water quality of a specific water body, which would significantly decrease the load of data for further data processing.

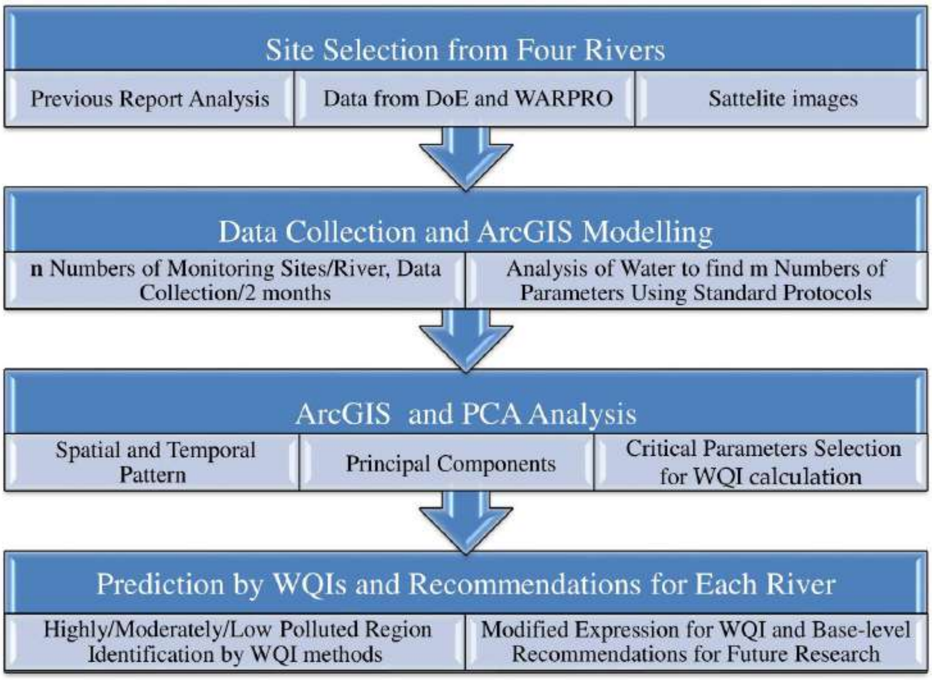
In this research, PCA would be utilized to reduce the huge amount of data into some meaningful, simpler data that can be further used to find out the prominent factors that are responsible for the variation of the water quality characteristics. To achieve this, MS Excel 2016 and IBM SPSS Statistics 28.0.0.0 (Armonk, NY, USA) would be used. SPSS (Statistical Package for the Social Sciences) is a software tool that is frequently utilized frequently for the purpose of statistical analysis in a variety of fields. It has numerous applications in dealing with data. IBM

SPSS is well-known for its applications, which include advanced statistical analysis, pre-processing, processing, and analysis of data. IBM SPSS is used in different fields such as stock market research, survey marketing, government organizations, information technology, etc [15]. The normalization and dimension reduction methods of SPSS would be utilized in this study. The input data were fed to the dimension reduction menu in SPSS for generating principal components. PCA would be conducted on the variables that had been normalized in order to obtain major principal components (PCs) and also to minimize the participation of variables that had a minor significance. These PCs were then exposed to varimax rotation (raw) in order to generate Vari-factors using Varimax with a Kaiser normalization method. Then, we can identify the factor within these PCs, those were significant for the water quality measurement.

After identifying the critical parameters, these would be used to calculated the WQI. As, proposing WQI for each rivers utilizing the bulk amount of information of several quality parameters is time consuming, and feasible in terms of economy and labor. Thus, if we can find the parameters that have the most influence in WQI, huge amount of data would be reduced. This can reduce the cost of the future projects concerning the river water quality analysis.

Water Quality Index:

The WQI is an index that measures the combined influence of numerous water quality variables [16]. The water quality index was established to analyze natural and artificial activities based on fundamental groundwater chemistry markers. The water quality index can also assess the quality of river water in general as well as in terms of its intended use: for drinking, for recreation, for aquatic zone, for agriculture, etc. [17]. There are many water quality index calculation methods available for determining the water quality such as National Sanitation Foundation WQI, Oregon WQI, Weighted Arithmetic WQI, and Canadian Council of Ministers of the Environment WQI [18]. We would calculate the WQI using the critical parameters obtained from PCA. The obtained WQIs from different models would be compared. Then, based on the research and PCA, we would propose a modified expression of WQI based on the National Sanitation Foundation WQI calculation for Bangladeshi rivers, which would make, the WQI calculations quick, objective, and reproducible and enable the evaluation of changes in water quality in various regions.
Overall flow chart of the methodology is shown as below:



EXPECTED OUTPUT:

The output of the research would cover the followings:

- Baseline database of water quality for Buriganga, Turag, Balu and Shitalakshya rivers
- The most critical water quality parameters for each selected river
- Standardized Water Quality Index (WQIs) for each selected rivers
- Recommendation for future river rehabilitation initiatives and effective monitoring system.

The outputs would be delivered in the form of reports. Moreover, the finding of this study would be published in peer-reviewed open-access journals and conference proceedings. The results of the research would be presented at national and international conferences/seminars for open discussion. All the datasets and results would also be archived in the National Water Resources Database (NWRD) of Bangladesh for dissemination to other research groups.

10

REFERENCES:

1. Bashar, T. and I.W.H. Fung, *Water Pollution in a Densely Populated Megapolis, Dhaka.* Water, 2020. **12**(8): p. 2124.
2. Kamal, M.M., A. Malmgren-Hansen, and A. Badruzzaman, *Assessment of pollution of the River Buriganga, Bangladesh, using a water quality model.* Water science and technology, 1999. **40**(2): p. 129-136.
3. Akbor, M., M. Uddin, and M. Ahsan, *Investigation of water quality parameters at different points in the Buriganga river, Bangladesh.* Journal of Environmental Science and Natural Resources, 2017. **10**(1): p. 75-80.
4. Whitehead, P., et al., *Restoring water quality in the polluted Turag-Tongi-Balu river system, Dhaka: Modelling nutrient and total coliform intervention strategies.* Science of the total environment, 2018. **631**: p. 223-232.
5. Kabir, A., et al., *Eco-environmental assessment of the Turag River in the megacity of Bangladesh.* Environmental Challenges, 2022. **6**: p. 100423.
6. Oseke, F.I., et al., *Assessment of water quality using GIS techniques and water quality index in reservoirs affected by water diversion.* Water-Energy Nexus, 2021. **4**: p. 25-34.
7. Oketola, A.A., S.M. Adekolurejo, and O. Osibanjo, *Water quality assessment of River Ogun using multivariate statistical techniques.* 2013.
8. Tharmar, E., et al., *Hydrogeochemistry and Water Quality Assessment in the Thamirabarani River Stretch by Applying GIS and PCA Techniques.* Sustainability, 2022. **14**(24): p. 16368.
9. Nandini, N., et al., *Application of Remote Sensing, GIS, and GPS Technology on Environment and Earth's Natural Resources Management An Introduction.* Mapana Journal of Sciences, 2007. **6**(1): p. 96-117.
10. Wu, C.-Y., et al., *Comparison of different spatial interpolation methods for historical hydrographic data of the lowermost Mississippi River.* Annals of GIS, 2019. **25**(2): p. 133-151.
11. Nath, B., et al., *GIS mapping-based impact assessment of groundwater contamination by arsenic and other heavy metal contaminants in the Brahmaputra River valley: a water quality assessment study.* Journal of Cleaner Production, 2018. **201**: p. 1001-1011.
12. Lu, G.Y. and D.W. Wong, *An adaptive inverse-distance weighting spatial interpolation technique.* Computers & geosciences, 2008. **34**(9): p. 1044-1055.
13. Penkova, T.G., *Principal component analysis and cluster analysis for evaluating the natural and anthropogenic territory safety.* Procedia computer science, 2017. **112**: p. 99-108.
14. Bengraïne, K. and T.F. Marhaba, *Using principal component analysis to monitor spatial and temporal changes in water quality.* Journal of Hazardous Materials, 2003. **100**(1-3): p. 179-195.
15. Tufféry, S., *Data mining and statistics for decision making*. 2011: John Wiley & Sons.

16. Pesce, S.F. and D.A. Wunderlin, *Use of water quality indices to verify the impact of Córdoba City (Argentina) on Suquía River.* Water research, 2000. **34**(11): p. 2915-2926.
17. Al-Shujairi, S.O.H., *Develop and apply water quality index to evaluate water quality of Tigris and Euphrates Rivers in Iraq.* International Journal of Modern Engineering Research, 2013. **3**(4): p. 2119-2126.
18. Zotou, I., V.A. Tsihrintzis, and G.D. Gikas, *Water quality evaluation of a lacustrine water body in the Mediterranean based on different water quality index (WQI) methodologies.* Journal of Environmental Science and Health, Part A, 2020. **55**(5): p. 537-548.

Terms of References

PROFESSIONAL INPUT, QUALIFICATION AND TASKS OF RESEARCH TEAM:

Professional Input:

The overall period for the research has been considered for **2 (two) years** commencing from the date of signing the contract. It is estimated that for carrying out the research including relevant data collection, the Research team of BUET will require about **36 person-months** for this research. The estimated staff requirements for the research have been given below in the following table:

Research Team

Sl. No.	Position	Proposed Man- month
1	Principal Investigator	12
2	Research Assistant (2 nos.)	24
	Total	**36**

13

Requisite Qualification, Experience and Tasks of Key Personnel:

Sl. No	Key Personnel Designation	Qualification and Requisite Experience	Tasks (His/her major responsibilities shall include but not necessarily be limited to the following)
1.	Principal Investigator (PI), Professor and Head, Department of Chemical Engineering, BUET, Dhaka-1000	PhD in Water Resources Engineering/Civil Engineering/ Environmental Engineering. Has more than 20 years of experience in industry, teaching and academic research. Has 60 publications on water and wastewater treatment in international journals with high impact factors (CV is attached).	• The overall responsibility is to coordinate the research team. • Shall maintain all external and internal contacts requisite in the interest of the research. • PI would contribute to an analysis of process description, contribute to determining approach/methods for framing, decision making conforming to delegated rules. • Analyzing and interpreting the model results. • Assisting the concerned professional in the preparation of reports and scientific papers. • Orient the work plan and necessary training program in consultation with the Research Coordinator for technology transfer. • Prepare time frame for publication of different reports and scientific articles. • Arrange and supervise field visits and site investigations in the study area. • Attending meetings and seminars as and when required for the research implementation. • Any and all other works necessary for proper progress of the study.

14

Sl. No	Key Personnel Designation	Qualification and Requisite Experience	Tasks (His/her major responsibilities shall include but not necessarily be limited to the following)
2.	Research Assistant (2 nos.)	He/she should have a minimum of a Bachelor degree in Civil Engineering/Chemical Engineering/Water Resources Engineering/ Environmental Engineering or similar other degrees.	• Site survey of 4 rivers and sample collection. • Analysis of water and sediment samples in the laboratory • Data management and data analysis relating to river water and sediment quality. • Assist to set up a numerical/statistical model and simulate different scenarios of river water and sediment qualities. • Application of the model for prediction of river and sediment quality. • Application of deep learning techniques • Any other assignment requested by Principal Investigators and Research Coordinator.

15

DURATION OF THE RESEARCH:

The overall period for the research has been considered for **2 (two) years** commencing from the date of signing the contract.

REPORTING AND WORKSHOP:

Reporting:

The following reports are required to submit to WARPO time to time as per the following schedule:

Sl. No.	Report	Deadline	Copies
1.	Inception Report	end of the 2nd Month	10 Copies
2.	Interim Report	end of the 12th Month	10 Copies
3.	Draft Final Report	end of the 21st Month	10 Copies
4.	Final Report	end of the 23rd Month	15 Copies

Major Deliverables:

The followings will be expected deliverable after completion of this research:

- All primary and secondary datasets collected from different organizations
- All developed models and predictive tools
- All published reports, conference and journal papers from the research

Total Cost:

The estimated cost of this research is **180.16 Lac BDT** (One Crore Eighty Lac Sixteen Thousand Taka only).

Workshops:

The following 2 (two) workshops will be organized as a part of consultation and dissemination of outcome of the research:

Sl. No.	Workshop	Tentative Schedule
1.	Inception Workshop	Soon after submission of the Inception report
2.	Final Workshop	Soon after submission of the Draft Final report

The first workshop will be a launching workshop, organized at Dhaka, after submission of the Inception Report. The workshop will be participated by the Government officials, experts and professionals working in the sector of river management. The main objective of the inception workshop is to share the objectives, methodologies and approaches to be followed in the research. The Final workshop will be a National workshop, organized after the submission of the Draft Final Report. The workshop will also be participated by Government high officials, International financial partners and reputed experts in water quality modeling. The objective of the workshop would be to disseminate the findings and incorporate valuable suggestions from different experts on the research carried out.

PROJECT ORGANIZATION:

The specialized professionals of WARPO shall assist the Research team as required for the research. WARPO team will be compromised of the following professionals:

Sl. No.	WARPO Professionals	Position
1.	Senior Scientific Officer (Navigation)	Principal Research Coordinator
2.	Senior Scientific Officer (Groundwater)	Research Associate
3.	Senior Scientific Officer (Water)	Research Associate
4.	Scientific Officer (Assistant Programmer)	Research Assistant
5.	Scientific Officer (Navigation)	Research Assistant

The Research team will work under the overall administrative responsibility and technical supervision of Water Resources Planning Organization (WARPO), led by the Research

Coordinator. The Research team will ensure that all activities/assignments will be executed in line with the scope of work and agreed work plan. In case changes/modifications to the agreed work plan, the principal investigator will bring this to the attention of WARPO.

For smooth and proper completion of the research, a Technical Committee, to be chaired by the Director General, WARPO will be formed. The Committee will be responsible for overall guidance of the research. Each committee member will get honorarium of Tk 2000.00 for every meeting. The TC will be formed comprising the following officials from different agencies:

Technical Committee (TC)		
1.	Director General, Water Resources Planning Organization (WARPO)	**Chairman**
2.	Director (Technical), Water Resources Planning Organization (WARPO)	Member
3.	Representative from Ministry of Water Resources (MoWR)	Member
4.	Representative from Department of Water Resource Engineering, BUET	Member
5.	Representative from Institute of Water and Flood Management, BUET	Member
6.	Representative from Chief Planning office, BWDB	Member
7.	Representative from Department of Environment (DoE)	Member
8.	Representative from Department of Civil Engineering, UAP	Member
9.	Principal Scientific Officer (Environment, Forests & Fisheries), WARPO	Member
10.	Principal Scientific Officer (Senior System Analyst), WARPO	Member
11.	Principal Scientific Officer (Water Resources), WARPO	Member
12.	Principal Scientific Officer (Monitoring), WARPO	Member
13.	Principal Scientific Officer (Engineering), WARPO	Member
14.	Principal Investigator, BUET	Member
15.	Principal Research Coordinator, WARPO	Member
16.	Research Associate (concerned), WARPO	Member-Secretary

Terms of References of the Technical Committee

- Providing necessary policy guidelines, institutional support and other directives for the successful implementation of the research.
- Facilitating coordination among the inter-agencies activities, monitoring and overall supervision of the research.
- The Committee will meet at least once in every 4(four) month, and in emergency cases as and when necessary.

18

TASKS AND RESPONSIBILITIES:

This a collaborative research between WARPO and BUET. The Research team will work together and jointly perform the tasks specified in the scope of work. However, the tasks can be assigned to WARPO and BUET separately as leading agency based on their areas of expertise. WARPO, being the policy developing organization, will take the lead on developing WQI framework, policy recommendation and disseminating research findings. BUET, being the technical university, can take the lead on technical aspects of the research such as water sample collection, laboratory analysis, PCA analysis and capacity building among stakeholders. Hence, the major task distribution between them has been listed below:

Tasks led by WARPO	Tasks led by BUET
a) Secondary data collection and analysis	a) Literature review of previous reports
b) Select potential monitoring sites	b) Water sample collection
c) Procure instrument and equipment	c) Sample storage and processing
d) Laboratory testing and analysis	d) Perform Principal Component Analysis (PCA) to identify critical parameters
e) Develop framework for calculating WQI	e) Publish research findings in academic journals
f) Analyze the WQI to identify trends across different seasons	f) Conduct training for capacity building and technology transfer
g) Develop recommendations for policy-makers and stakeholders	g) Submission of reports
h) Disseminate research findings through workshop/seminar	
i) Maintain the time schedule/work plan	

19

WARPO's Responsibilities:

The Research Coordinator will ensure that the objectives of the research as detailed in the Terms of References (ToR) are achieved within the agreed time schedule. He will in the context of the ToR direct the research process and supervise the execution of the research and monitor progress according to the said objectives. In particular, the Research Coordinator will take necessary action in good time where such monitoring shows that outputs are not likely to be supplied at the required time. The BUET Research team will be required to reschedule activities on the direction of the Research Coordinator if this becomes necessary.

The BUET Research team will have regular meetings with WARPO professionals to discuss technical and management issues. Any unresolved issue should be taken up with the Director General, WARPO. WARPO team will make arrangements for the BUET Research team to meet the concerned GoB agencies in Dhaka to enable the agencies to be aware of the research objectives from its inception, particularly in relation to such matters as data collections, stakeholder's consultation etc.

WARPO will be responsible for arranging the following facilities:

- Previous reports on river water quality that are available with WARPO.
- Provide assistance for arranging the collection of data from NWRD and other agencies, if needed.
- Make available WARPO professionals to assist site investigation for sampling point selection as required for the research.
- Make available information from other study components.

BUET's Responsibilities:

The BUET Research team shall carry out the services as detailed in the 'Scope of Work' in the best interest of the research with reasonable care, skill and diligence with sound engineering, administrative and financial practices. The Principal Investigator will prepare time frame for publication of different reports and scientific articles through consultation with the Research Coordinator. He will also be responsible to the Director General, WARPO for proper and timely execution of all the activities of the research mentioned in the ToR of the research project.

20

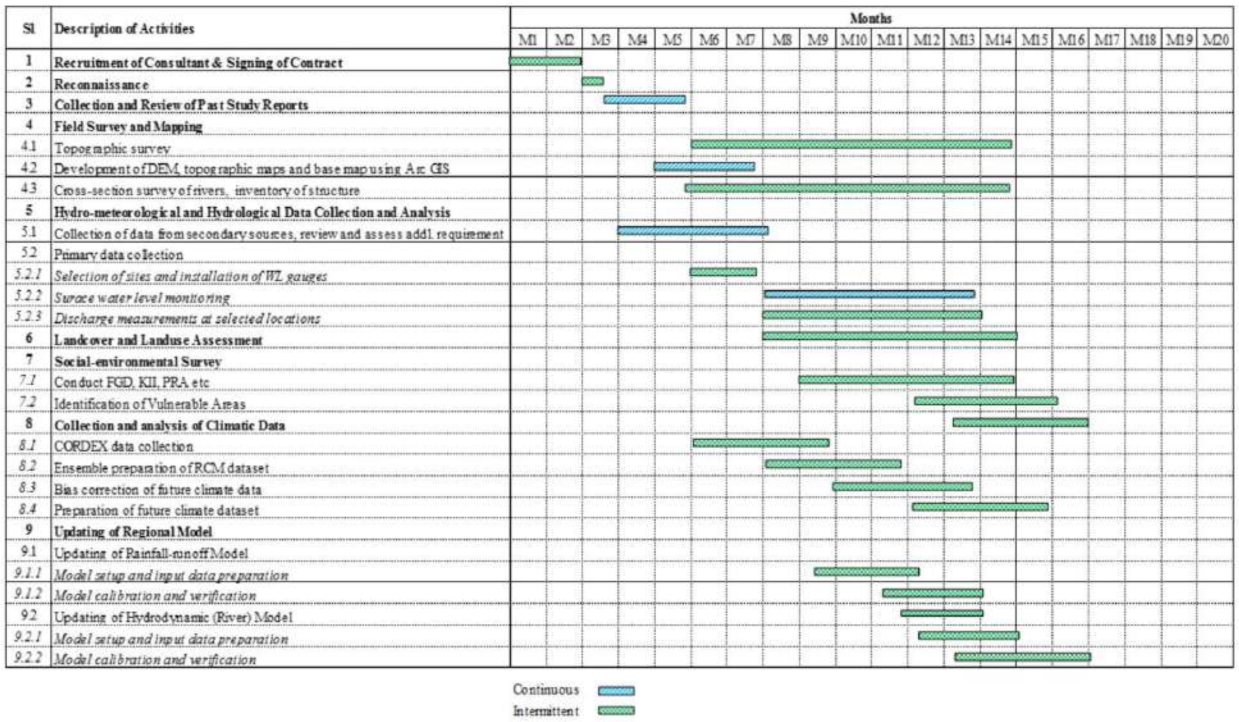
The BUET Research team will be responsible for arranging the following facilities:

- Make necessary arrangements for site investigation and sample collection as required for the research.
- Making necessary arrangements for additional data collection from secondary sources as needed for the research.
- Facilitate the laboratory testing of the collected samples.
- Carrying out activities as per the scope of work and delivering the research reports and progress reports at regular intervals.
- Handing over the research results to WARPO for their use and records.
- Arranging necessary training for WARPO professionals as a matter of technology transfer.

21

WORK PLAN:

Total time required for the research would be **2 (two) years.** The timeline for the different activities are shown below:



22

"Research on Establishment of Water Quality Index (WQI) through Principal Component Analysis for the Dhaka-based Rivers" শীর্ষক গবেষণা প্রস্তাবের উপর পানি সম্পদ মন্ত্রণালয়ের গবেষণা ব্যবস্থাপনা কমিটির সভার সিদ্ধান্ত প্রতিপালন

গবেষণা ব্যবস্থাপনা কমিটির সভার সিদ্ধান্ত	প্রতিপালন অগ্রগতি
১.১ Research Hypothesis যথাযথভাবে উল্লেখপূর্বক Research Proposal পরিমার্জন করার সিদ্ধান্ত গৃহীত হয়।	Research Proposal এ যথাযথভাবে Research Hypothesis উল্লেখ করে (পৃষ্ঠা নং - ৫) Research Proposal পরিমার্জন করা হয়েছে।
১.২ Household, small industry, Heavy industry (RMG, Tannery, Plastic) নিকটবর্তী এলাকায় দূষণ চিত্র তুলে ধরার সিদ্ধান্ত গৃহীত হয়।	পানির নমুনা সংগ্রহের জন্য Household, small industry, Heavy industry (RMG, Tannery, Plastic) এর নিকটবর্তী এলাকাকে অগ্রাধিকার দিয়ে প্রতিটি নদীতে ৬টি করে উপযুক্ত মনিটরিং সাইট নির্বাচন করা হবে।
১.৩ ToR নির্ধারন এবং তদানুসারে সুস্পষ্টভাবে কার্যক্রমের বিবরণ, কর্মবন্টন উল্লেখ করার সিদ্ধান্ত গৃহীত হয়।	গবেষণা কার্যক্রমের বিবরণ, ওয়ারপো ও বুয়েট টীমের কর্মবন্টন (পৃষ্ঠা নং - ১৯), দায়িত্ব, রিপোর্টিং সুস্পষ্টভাবে উল্লেখপূর্বক ToR নির্ধারন করা হয়েছে (পৃষ্ঠা নং ১৩-২১)।
১.৪ বাংলাদেশের ঋতুগত Seasonal Variation (গ্রীষ্মকাল, বর্ষাকাল, শীতকাল) এবং সাংস্কৃতিক (ঈদ, পূজা) কারনে পানিতে দূষণের মাত্রাগত পরিবর্তনের চিত্র নির্ণয় করার সিদ্ধান্ত গৃহীত হয়।	বাংলাদেশের ঋতুগত Seasonal Variation এর দরুন পানিতে দূষণের মাত্রাগত পরিবর্তনের প্রকৃত চিত্র নির্ণয় করার জন্য pre-monsoon, monsoon and dry season সময়ে পানির নমুনা সংগ্রহ করা হবে এছাড়া সাংস্কৃতিক উৎসব (ঈদ, পূজা) সময়কালীন নদীগুলো থেকেও পানির নমুনা সংগ্রহ করা হবে। ডাটা কালেকশন অংশে যথাযথভাবে এর উল্লেখ করা হয়েছে (পৃষ্ঠা নং - ৬)।
১.৫ গবেষণা বাজেট ও সময়কাল ২৪ মাস পর্যন্ত বর্ধিতকরণের মাধ্যমে Research Proposal পরিমার্জন করার সিদ্ধান্ত গৃহীত হয়।	গবেষণার কার্যক্রম আরো ৬ মাস বৃদ্ধি করে মেয়াদকাল সর্বমোট ২৪ মাস পর্যন্ত বর্ধিত করে Research Proposal পরিমার্জন করা হয়েছে। পরিমার্জিত কর্মপরিকল্পনা (পৃষ্ঠা নং - ২২) এবং পুনর্গঠিত বাজেট (পৃষ্ঠা নং - ২৩) প্রস্তাবে সংযোজন করা হয়েছে।
১.৬ পরিমার্জনকৃত Research Proposal আগামী ১০ মে ২০২৩ তারিখের মধ্যে পানি সম্পদ মন্ত্রণালয়ে প্রেরণের সিদ্ধান্ত গৃহীত হয়।	গবেষোনা কমিটির সভার সিদ্ধান্ত/সুপারিশ অনুসরণপূর্বক পরিমার্জনকৃত Research Proposal পানি সম্পদ মন্ত্রণালয়ে পাঠানো হয়েছে।

26

Appendix J: Contract

৳১০০ ৳১০০

একশত টাকা

খন ৫৪৭৫৪৫১

MEMORANDUM OF AGREEMENT

A COLLABORATIVE RESEARCH PROJECT TITLED

"Establishment of Water Quality Index (WQI) through Principal Component Analysis for the Dhaka-based Rivers"

Between

Water Resources Planning Organization (WARPO)
Ministry of Water Resources
Government of Bangladesh

And

Research and Innovation Centre for Science and Engineering (RISE)
Represented by the Department of Chemical Engineering (ChE)
Bangladesh University of Engineering and Technology (BUET)
Dhaka, Bangladesh

This Memorandum of Agreement (MoA) is signed on the 06 day of June 2023 between Water Resources Planning Organization, Ministry of Water Resources, Government of Bangladesh (hereinafter referred to as '**WARPO**') and Research and Innovation Centre for Science and Engineering (RISE), represented by the Department of Chemical Engineering, Bangladesh University of Engineering and Technology (hereinafter referred to as '**BUET**'). WARPO and BUET both recognizing the need for collaboration entered into this MoA based on the following agreement in order to conduct the above mentioned research project:

1. RESEARCH WORK:

a) Select potential monitoring sites for each river through field survey and GIS technique.
b) Collect river water samples from each monitoring site following the standard protocols, covering pre-monsoon, monsoon and dry seasons.
c) Perform laboratory tests and analysis of water quality parameters e.g., pH, TDS, DO, BOD, COD, EC, Total Coliform etc. following standard guidelines.
d) Evaluate Principal Component Analysis (PCA) to find critical components (PCs) for each site utilizing statistical data regression.
e) Derive an appropriate method to calculate WQI for the selected rivers following internationally adopted WQI approach.
f) Recommend and assist the policy makers with a possible route for a river rehabilitation.
g) Arrange seminars on the research findings and training programs on river quality assessment and monitoring techniques.
h) The basic methodology and approach of the 2-year collaborative research project is given in Annexure-I.
i) The Principal Research Coordinator from WARPO and the Principal Investigator from BUET will be responsible for research planning and activities.

"দেশপ্রেমের শপথ নিন, দুর্নীতিকে বিদায় দিন"

MoA 1

৳১০০ ৳১০০

একশত টাকা

খন ৫৪৭৫৪৫০

2. MAJOR DELIVERABLES:

a) BUET will deliver all the required reports to WARPO on time (as mentioned in Annexure-I).
b) WARPO and BUET will jointly publish the outcome of this collaborative research project in international peer reviewed journals.
c) Upon completion of the research, BUET will hand over all primary and secondary datasets to WARPO.

3. DURATION OF THE RESEARCH:

The overall period for this research project is **2 (two) years**, commencing from 21 June 2023 and ends on 20 June 2025. The project duration can be extended, if necessary, upon agreement between two parties.

4. CONTRACT PRICE WITH BUET:

Total estimated budget of this research is 180.16 Lac BDT (one crore eighty lac sixteen thousand taka only). Within this budget, the contract price with BUET is **65.91 Lac BDT** (sixty-five lacs ninety-one thousand taka only). The contract price includes remuneration of the Principal Investigator and Research Assistants, field survey, sample collection and training as described in **Annexure-I.**

5. MODE OF PAYMENT:

The mode of payment will be as follows:

Instalment No.	Description of Work	Payment (%)
1.	Preparatory Activities on Signing of Agreement	25% of the Contract Price
2.	On Submission of Inception Report	15% of the Contract Price
3.	On Submission of Interim Report	40% of the Contract Price
4.	On Submission of Final Report	20% of the Contract Price

All payments shall have to be made through account payee cheques or bank transfer to the bank account detailed below. VAT and Tax will be deducted at source, the copies of challans of deposited VAT & Tax shall have to be submitted to RISE office along with the payment.

Name of Bank	:	Sonali Bank Limited
Bank Address	:	BUET Branch
Name of Account	:	RISE Centre, BUET
Account Number	:	4404002000852

"দেশপ্রেমের শপথ নিন, দুর্নীতিকে বিদায় দিন"

MoA_2

৳১০০ ৳১০০

একশত টাকা

খন ৫৪৭৫৪৪৯

6. **VALIDITY:** This MoA is valid for a period of 25 (twenty-five) months. It can be renewed, if necessary, upon agreement between the parties.

7. **AMENDMENT AND MODIFICATION:** Either party may request in writing an amendment or modification of this MoA. Any amendment or modification agreed by the parties, shall be produced in writings and shall form a part of this MoA.

8. **CONFIDENTIALITY:** Both the parties during and after the end of the MoA must not disclose any such matter to anybody not part of this MoA, which has come to their knowledge during the course of their actions following this MoA.

9. **SETTLEMENT OF DISPUTE:** Any dispute or difference between the parties concerning this MoA shall be settled amicably through mutual consultation and/or negotiations between the parties.

10. **GOVERNING LAW:** This MoA is subject to the laws of The People's Republic of Bangladesh.

IN WITNESS WHEREOF both the parties have executed the Memorandum of Agreement (MoA) as the date, month and year written above.

Md. Rezaul Maksud Jahedi Director General Water Resources Planning Organization Ministry of Water Resources 72 Green Road, Dhaka-1215	Dr. Muhammad Anisuzzaman Talukder Professor and Director Research and Innovation Centre for Science and Engineering Bangladesh University of Engineering and Technology (BUET), Dhaka-1000
Witness: Kazi Saidur Rahman Senior Scientific Officer Water Resources Planning Organization Ministry of Water Resources 72 Green Road, Dhaka-1215	Dr. Md. Shahinoor Islam Professor and Head Department of Chemical Engineering Bangladesh University of Engineering and Technology (BUET), Dhaka-1000

"দেশপ্রেমের শপথ নিন, দুর্নীতিকে বিদায় দিন"

MoA_3

Appendix K: Selected Sites for Sample Collection

Site Selection was done with a prior site visit and inspection. A total 36 sites were selected for sample collection with 9 sites for each river. Following is a summary table for selected sampling sites with their coordinates.

Table 12.2: Geographical location of selected sites

River Name	Site Code	Sampling Sites	Coordinates	
			Latitude	Longitude
Buriganga	BG1	*Gabtoli, Dhaka*	23°46'49.0"N	90°20'13.6"E
	BG2	*Basila, Dhaka*	23°44'58.6"N	90°20'21.6"E
	BG3	*Gudaraghat, Hazaribag*	23°44'02.1"N	90°21'13.9"E
	BG4	*Kholamora Ghat, Kamranghirchar*	23°42'53.0"N	90°21'34.9"E
	BG5	*Chandir Ghat*	23°42'39.8"N	90°23'26.5"E
	BG6	*Sadarghat Launch Terminal*	23°42'21.6"N	90°24'20.8"E
	BG7	*Postogola Bridge, Dhaka*	23°41'16.0"N	90°25'37.8"E
	BG8	*Pagla Ghat, Dhaka*	23°39'40.3"N	90°27'15.1"E
	BG9	*Fatullah Lauch Terminal*	23°38'24.8"N	90°28'21.0"E
Turag	TR1	*Kashimpur, Namabazar*	23°59'03.6"N	90°19'40.0"E
	TR2	*Mohisher Tek*	23°57'02.8"N	90°20'41.9"E
	TR3	*Ashulia Kacha Bazar*	23°53'48.0"N	90°20'03.3"E
	TR4	*Near Fulpukuria Thread, Tongi*	23°53'51.6"N	90°26'05.4"E
	TR5	*Tongi Rail Bridge, Gazipur*	23°52'55.4"N	90°24'20.3"E
	TR6	*Ashulia Landing Station*	23°53'29.5"N	90°21'34.7"E
	TR7	*Rustompur Ghat, Mirpur road*	23°52'29.2"N	90°21'00.8"E
	TR8	*Birulia*	23°50'50.0"N	90°20'13.9"E
	TR9	*Diabari Ghat, Mirpur road*	23°47'53.9"N	90°20'24.2"E
Balu	BL1	*Pubail Bridge*	23°56'04.6"N	90°29'23.5"E
	BL2	*Ulukhola Bazar*	23°54'01.4"N	90°28'56.3"E
	BL3	*Talia Termukh Bridge*	23°52'55.4"N	90°27'38.4"E
	BL4	*Near Balu Bridge, 300 feet, Dhaka*	23°50'04.6"N	90°28'41.7"E
	BL5	*Near LGED Bridge, Jolshiri*	23°48'58.5"N	90°29'09.9"E
	BL6	*Beraid Boat Ghat, Beraid*	23°48'05.9"N	90°28'53.9"E
	BL7	*Near Edarkandi-Fakirkhali Road*	23°46'36.9"N	90°28'43.4"E
	BL8	*Near Eastern Straw Board*	23°45'23.8"N	90°29'11.1"E
	BL9	*Rajakhali Ghat, Demra*	23°44'07.3"N	90°29'28.3"E
Sitalakhya	SL1	*Kholapara-Ghorashal Kheyaghat*	23°56'27.4"N	90°36'57.4"E
	SL2	*Kaliganj Kazir Char Kheyar Ghat*	23°54'57.1"N	90°34'01.9"E
	SL3	*Beldi Bazar, Beldi*	23°52'01.6"N	90°33'17.2"E
	SL4	*Near Hatabo Bazar, Rupganj*	23°48'33.6"N	90°32'37.2"E
	SL5	*Habib Ghat, Boralu Bazar*	23°45'46.8"N	90°30'28.0"E
	SL6	*South Rupshi Masjid Ghat, Rupshi*	23°43'55.1"N	90°30’30.8"E
	SL7	*Horipur Power Plant*	23°40'57.3"N	90°31’46.9"E
	SL8	*Nabjiganj Ghat*	23°37'59.0"N	90°31’05.0"E
	SL9	*BIMT Launch Ghat*	23°36'26.9"N	90°30’21.0"E

Following are the pictures taken in inspection visits for site selection -

Surrounding in BG2 Site


Surrounding in BG3 Site


Surrounding in BG4 Site


Surrounding in BG5 Site


Surrounding in BG6 Site


Surrounding in BG7 Site


Surrounding in BG8 Site


Surrounding in BG9 Site


Surrounding in TR1 Site


Surrounding in TR2 Site


Surrounding in TR3 Site


Surrounding in TR4 Site


Surrounding in TR5 Site


Surrounding in TR6 Site


Surrounding in TR7 Site


Surrounding in TR8 Site


Surrounding in TR9 Site


Surrounding in BL1 Site


Surrounding in BL2 Site

Surrounding in BL3 Site

Surrounding in BL4 Site

Surrounding in BL5 Site

Surrounding in BL6 Site

Surrounding in BL7 Site

Surrounding in SL1 Site


Surrounding in SL2 Site


Surrounding in SL3 Site


Surrounding in SL4 Site


Surrounding in SL5 Site


Surrounding in SL6 Site


Surrounding in SL7 Site


Surrounding in SL8 Site


Surrounding in SL9 Site


Appendix L: Field Visit Images


Field visit at Buriganga

Field visit at Turag

Field visit at Balu

Field visit at Shitalakhya

Appendix M: Sample Collection Images

Figure 12.15: Sample collection images from Buriganga River

Figure 12.16: Sample collection images from Balu River

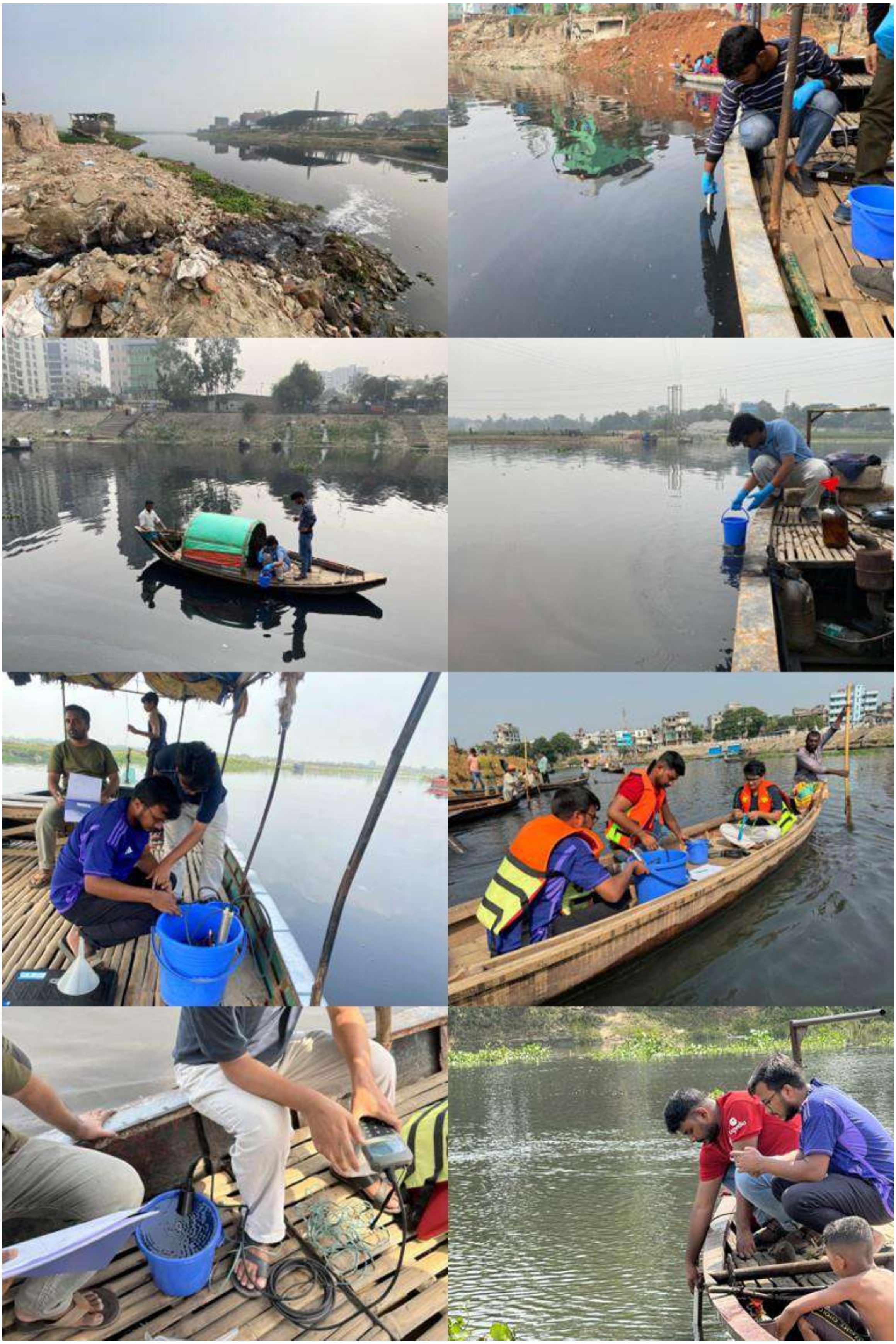


Figure 12.17: Sample collection images from Turag River

Figure 12.18: Sample collection images from Sitalakhya River