SURFACE WATER QUALITY ASSESSMENT BY ENVIRONMETRICS APPROACH

A thesis submitted

in partial fulfillment of the requirement

for the degree of

Doctor of Philosophy

Submitted

by

KUNWAR RAGHVENDRA SINGH

Department of Civil Engineering

Indian Institute of Technology Guwahati

Guwahati -781039, Assam, India

May, 2019

Indian Institute of Technology Guwahati

Assam, India

Candidate's Declaration

This is to certify that the thesis entitled "SURFACE WATER QUALITY ASSESSMENT BY ENVIRONMETRICS APPROACH", submitted by me to the Indian Institute of Technology Guwahati, for the award of the degree of Doctor of Philosophy, is a bonafide work carried out by me. The content of this thesis, in full or in parts, have not been submitted to any other University or Institute for the award of any degree.

Date:

Kunwar Raghvendra Singh
Registration no. 156104005
Department of Civil Engineering
Indian Institute of Technology Guwahati
Guwahati-781039, Assam, India.

Indian Institute of Technology Guwahati

Assam, India

Certificate

This is to certify that the thesis entitled "SURFACE WATER QUALITY ASSESSMENT BY ENVIRONMETRICS APPROACH", submitted by Kunwar Raghvendra Singh (156104005), a research scholar in the Department of Civil Engineering, Indian Institute of Technology Guwahati, for the award of the degree of Doctor of Philosophy, is a record of an original research work carried out by him under our supervision and guidance. The thesis has fulfilled all requirements as per the regulations of the institute and in our opinion has reached the standard needed for submission. The results embodied in this thesis have not been submitted to any other University or Institute for the award of any degree or diploma.

Dr. Ajay Kalamdhad

Associate Professor
Department of Civil Engineering
IIT Guwahati
Assam, India

Dr. Bimlesh Kumar

Associate Professor
Department of Civil Engineering
IIT Guwahati
Assam, India

Acknowledgements

First and foremost, my humble gratitude goes to my supervisors Dr. Ajay Kalamdhad & Dr. Bimlesh Kumar for their support and guidance all the time. This work would not have been completed without their valuable ideas, suggestions, and inputs. Their high expectations and endless dedication to this project have played a major role in the successful completion of this ambitious body of work. The thrust they provided to all his students for in time completion of their works by no means a meagre contribution.

I express my sincere thanks to members of my Doctoral Committee viz. Professor Saswati Chakraborti, Professor Kishore Nanda, Dr. Manish Kumar Goyal and Dr. Rishikesh Bharti, for their valuable suggestions and encouragement throughout my research work. I show my profound gratitude and respect to the former Head of the Civil Engineering Department, Professor Subashisa Dutta and present Head, Professor Chandan Mahanta for providing facilities for conducting the research.

It was my pleasure to associate with Dr. Waikhom Roshan Singh, Scientist, Manipur Pollution Control Board, Imphal during his days at Indian Institute of Technology Guwahati. The rounds of discussion we had in the field show the light of the day. A very special thanks goes to Dr. K. Dhamodharan, Dr. Arvind Kumar Shakya, Dr. V. Sudharsan Varma and Mr. Sachin Kumar Tomar for their support and encouragement throughout my research work. I am very thankful to Mr. Ankit Goswami and Mr. Rahul Dutta for assisting me during my experimental work and manuscript writing. I am also very thankful to Mr. Siddhant Das and Mr. Smitom for their valuable suggestions and helps during my thesis writing. I would be failing my duties, if I do not mention the role played by Mr. Sunil Mohan Dr. Gaurav, Mayur, Rohini, Getu Yoka and Venktesh, for their love and respect during my research work.

I greatly acknowledge the help and support that I received from my friends, colleagues, faculty and staff at IIT Guwahati. I would like to thank all Waste Management Research Group (WMRG) members who made my time unforgettable at IITG.

I am very thankful to the scientific officer of the department Ms. Jonali Saikia for extending all possible support. I gratefully acknowledge the unstinted help provided by Mr. Chitta Ranjan Medhi and Mr. Payodhar Pathak during all phases of work in environmental laboratory.

I wish to put on record the contributions of the respective Editors and Reviewers of all my published manuscripts through their valuable comments in upgrading my research study.

I would like to acknowledge the Ministry of Human Resource Development (MHRD) for the scholarship to carry out the doctoral thesis work.

My heartfelt thanks to parents, family members, and friends for moral support and encouragement to accomplish this endeavor.

Above all, I would like to thank the almighty for blessing me with wisdom, health and strength to accomplish the research work.

Kunwar Raghvendra Singh

ABSTRACT

Adulteration of surface water quality as a consequence of natural or anthropogenic activities has garnered global attention for their conservation and protection. Discharges of untreated municipal and industrial wastes, agricultural run-offs, leaching from landfills, mining and other activities have not only been hostile to aquatic ecosystems but also have introduced various trace elements such as toxic metals. Extensive programs for water quality monitoring and assessment have thus surfaced worldwide so as to have a proper insight on activities causing the degradation of such resources. These monitoring programs have become quintessential for developing a keen understanding on water quality processes for decision makers to comprehend, interpret and utilize this information in developing conservation strategies for the water resource. However, a global challenge has emerged in handling large sets of random data generated in these monitoring programs and utilizing them to derive useful information about the water quality of the resource. Environmetrics methods prove to be effective tools for interpreting large and complex water quality datasets as well as assessing the of spatio-temporal variations in surface water quality.

In the present study, various statistical methods were utilized as tools for water quality assessment of Brahmaputra River and its seven tributaries (Baralia, Puthimari, Pagladia, Beki, Manas, Kolong and Kameng River) as well as Deepor Beel, Assam (India). The study has been carried out in four phases. First phase includes the survey of study area as well as the collection and analysis of water samples. In the second phase, cluster analysis (CA), discriminant analysis (DA) and principal component analysis (PCA) were applied on the observed water quality data-sets. CA grouped all the sampling sites into three clusters, i.e. relatively less polluted (LP), medium polluted (MP), and highly polluted (HP) sites based on the similarities of the characteristics they possess. Result from CA was verified using DA, which helped in determining the continuous variables that discriminate two or more naturally occurring groups. Stepwise mode of DA was used which provided five (Na^+, Ca^{2+}, Mg^{2+}, SO_4^{2-} and Pb) and eight (DO, total alkalinity, K^+, Ca^{2+}, Mg^{2+}, Cl^-, SO_4^{2-} and Mn) water quality parameters affording 77.8 and 100% correct assignations in spatial analysis of Brahmaputra River and its tributaries respectively. It revealed the best results for spatial variation and provided an important data reduction. PCA applied to the three separate datasets obtained from CA resulted in six, five and five latent factors explaining 77.9, 91.68 and 79.29% of the total variance in the water quality datasets of LP, MP and HP sampling sites

respectively. In the third phase, Information entropy was used for developing a general water quality index based on the drinking water quality standards, heavy metal contamination index and an irrigation water quality index. A multi-criteria decision making method (TOPSIS) was also made use for overall ranking of sampling sites on the basis of entropy weights. Fourth phase was associated with the investigation of water quality variability and identification of ideal monitoring locations using entropy based diversity and disorder indices. Geospatial analysis and isoinformation lines were employed to generate geospatial maps that clearly depict the ideal monitoring locations which encountered highest variability in their water quality over the monitoring period with respect to physicochemical parameters; BOD, COD and heavy metals. Study of this phase introduced an innovative approach to derive maximum useful information by pin-pointing sampling stations for regular monitoring purposes.

Therefore, present study illustrates the necessity and usefulness of environmetrics methods for analysis and interpretation of large complex datasets with a view to get better information about the water quality and design of monitoring networks for effective management of water resources.

Keywords: Water quality, Statistical techniques, Shannon entropy, Heavy metals, TOPSIS, Diversity index, Disorder index

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

APHA	American Public Health Association
MSTs	Multivariate statistical technique
PCA	Principal component analysis
FA	Factor analysis
CA	Cluster analysis
DA	Discriminant analysis
WQI	Water quality index
EWQI	Entropy weighted water quality index
PI	Permeability index
KR	Kelley's ratio
MgR	Magnesium adsorption ratio
SAR	Sodium adsorption ratio
RSC	Residual sodium carbonate
ADD	Average daily dose
R_fD	Reference dose
HQ	Hazard Quotient
HI	Hazard Index
CI	Contamination Index
HPI	Heavy metal pollution index
HEI	heavy metal evaluation index
MCDMs	Multi-criteria decision making methods
TOPSIS	Technique for order of preference by similarity to ideal solution
POME	Principle of maximum entropy
DI	Disorder index

* All chemical symbols have their usual notation in the thesis

1

Chapter

Introduction

This chapter covers

- Introduction
- Need of the study
- Objectives of the study
- Scope of the study
- Thesis organization

1.1 INTRODUCTION

Innocuous and adequate amount of freshwater is essential for socio-economic development and smooth functioning of the ecosystems (Matta et al., 2017). Rivers and lakes are important sources of freshwater, crucial for the survival of civilizations, as they provide reliable water for our domestic usages and for various activities in agriculture, transportation and industry. Many civilizations flourished on the banks of rivers as the river water provided industrial, agricultural and economic prosperity (Bu et al., 2010; Tyagi et al., 2013; Li et al., 2016). In addition to providing goods and services to the local community in the form of staple food plants, inland fisheries and flood control (Maltby, 2013), lakes and other wetlands play a significant role in the environment principally acting as carbon sinks, water purifiers, host to biologically diverse ecosystems, etc. and contribute highly to the recharge of groundwater aquifers (Winter, 1999). Due to their socio-economic and ecological importance, freshwater resources lay the foundation for the growth and progress of a nation. As the global population increases, the demands on the world's freshwater supplies continue to grow. Hence, effective and sustainable management of water resources is vital for ensuring overall sustainable development.

Rampant urbanization, settlement and resettlement of communities aided with rapid industrialization has an adverse effect on water quality (Simeonov et al., 2003; Singh et al., 2004; Zhang et al., 2009; Wang et al., 2013). Surface water sources are now becoming a pathway through which wide varieties of biotic organisms are exposed to harmful elements either of anthropogenic or geological origin. Discharge of untreated or partially treated municipal and industrial wastewater, agricultural run-offs, leachates from landfills, and effluents from mining and other activities reaching the water bodies have not only been hostile to aquatic ecosystems, but also have introduced various trace elements such as toxic metals (Sin et al., 2001; Sundaray et al., 2006; Armitage et. al, 2007; Yuan et al., 2011). Existence of heavy metals in the form of ions, complexes, organically bound fractions and particulates in the water bodies or in bed sediments sheds light on their persistent, stable and non-degradable character (Pertsemli and Voutsa, 2007). In recent years, the water pollution of rivers, lakes and wetlands in most parts of the world as a consequence of various natural or anthropogenic activities has gained worldwide attention for their conservation and protection (Bastami et al., 2014; Fu et al., 2014). Thus, comprehensive and accurate assessments of trends in water quality as well as source identification of pollutants has emerged as the need of the hour. Consistent monitoring data becomes highly essential for such assessments and for evolving effective monitoring strategies.

Monitoring of surface water sources is essential for generating reliable information on its water quality which in turn will go a long way in preventing and controlling its pollution (Singh et al., 2004; Varol et al., 2012). A major problem in water quality monitoring is handling of huge and complex data sets generated due to the large number of water quality variables at different monitoring stations (Dixon and Chiswell, 1996). Elucidation of monitored data is a challenging task for researchers working in the field of surface water quality monitoring. Environmetrics approach, which provides quantitative environmental analysis, offers a better understanding of the observed data. In recent years, cluster analysis (CA), discriminant analysis (DA) and principal component analysis (PCA)/ factor analysis (FA) have been widely used for the interpretation of complex sets of water quality data and assessment of pollution sources (Shrestha and Kazama, 2007; Fan et al. 2010; Noori et al., 2010; Shihab and Abdul Baqi 2010; Batayneh and Zumlot 2012; Bouguerne et al., 2017). CA is an effective tool for determining the structure within complex datasets and for exploring similarities and/or dissimilarities among data and variables (Singh et al. 2004; Tsai et al. 2016). DA is used to determine continuous variables which discriminate between two or more naturally occurring groups. PCA is a pattern recognition technique that attempts to interpret the variance within a large set of inter-correlated variables by converting them into a smaller set of independent variables (Simeonov et al. 2003; Iscen et al. 2008). It provides information on the most significant parameters to describe the entire data set with minimum loss of original information (Iscen et al. 2008). These tools are effective means of manipulating, interpreting and representing data concerning surface water pollutants, geochemistry and sources of pollutants.

Evaluation of water quality from the monitored data is another important issue (Ongley, 1998). Monitoring programs evaluate a broad range of physical, chemical and biological water quality parameters as well as heavy metals. This necessitates their integration into simpler numerical scores which can be interpreted by trained experts as well as the general public. Most of these physico-chemical and biological parameters are integrated by Water Quality Indices (WQIs) into a single numerical score capable of describing the water quality at a particular location and at a particular point of time (Kaurish and Younos, 2007). Similar types of indices have been developed for metals as well. Heavy metal pollution indices (HPIs) have been used by various researchers to evaluate the heavy metal contamination in ground as well as surface waters.

Another important aspect of monitoring programs is prioritizing decisions to enforce policies and stream restoration efforts. In such circumstances, multi-criteria decision making methods (MCDMs) are widely considered by researchers in redressing management issues such as demand response, performance assessment, storage system, and renewable energy sources (Sianaki and Masoum, 2013). The efficiency of MCDMs have been reported regularly in the fields of water quality monitoring and assessment (Hernandez and Uddameri, 2013; Khan and Maity, 2016), and health risk assessment (Li et al., 2016).

Monitoring programs are also necessary for pinpointing trends in water bodies over a temporal scale as well as for efficient design of water quality monitoring networks (Ozkul et al., 2000). In fact, optimizing information needs and information gathered at low cost is one of the crucial aspects of monitoring programs (Ozkul et al., 2000). Most professionals engaged in the design and operation of the monitoring programs are familiar with the symptoms of "data rich but information poor" monitoring systems which generate large amount of data in a discrete form but are often incapable of describing water quality trends in an area (Ward et al., 1986).

It is essential to remove the significant gap between the information needs on water quality and the information produced by the monitoring systems. In recent years, several researchers have focused on the application of "information entropy" in water quality monitoring to derive maximum useful information at optimum costs. However, most of its potential still remains untapped in addressing the important aspects of monitoring (Singh., 2011). Information entropy developed by Claude Shannon in 1948 is capable of measuring the disorder or randomness associated with the occurrence of an event (Singh, 2013). It is a measure of the unpredictability of a random event, or equivalently the average information derived from its occurrence. In recent years, information entropy has had diverse applications in the field of hydraulic engineering and environmental engineering as a large number of random processes predominate our environment. In the light of information entropy, several techniques such as development of entropy based quality indices to address shortcomings of conventional WQIs, modification of conflicts between different WQIs by entropy based MCDMs, and optimization of on-line monitoring systems to refine the information gathered from monitoring programs have emerged.

Success of any good environmental policy depends on sound environmental monitoring (Jeong et al., 2006). Reliable information on water quality and identification of pollution sources are essential for prevention and control of surface water pollution (Simeonov et al., 2003; Shrestha and Kazama, 2007; Bu et al., 2010). However, there is significant gap between information required and the information produced by current water quality monitoring systems (Ozkul et al., 2000). It is essential to reassess the various components of water quality monitoring and explore and develop methodologies for optimizing between information needs and information gathered in effective and efficient way.

1.2 OBJECTIVES OF PRESENT STUDY

The overall aim of the present study is to identify the best possible approach for water quality monitoring program and the assessment of observed data sets in order to provide effective and efficient information at optimum cost. The following objectives are defined to focus on the primary aim of this study:

(a) Identification of latent pollution sources in water bodies.

(b) To explore and develop water quality indices for different purposes (i.e. drinking, irrigation) using information entropy.

(c) Assessment and application of a method for overall ranking of sampling sites in water bodies.

(d) To explore and develop a suitable methodology for identification of ideal sampling sites.

1.3 NEED OF THE STUDY

As per the report of World Health Organization (WHO), in 2015, more than 800 million people lack even a basic drinking-water service, including 159 million people who are dependent on surface water. Furthermore, the report suggested that by 2025, half of the world's population will be living in well stressed areas. Assam, a north-eastern state of India, though blessed with innumerable freshwater resources, is today facing a similar situation wherein there has been a scarcity in safe and potable water. The state is known for its tea plantation, petroleum resources, muga silk and bio-diversity. With a large portion of the state's economy based on agricultural activities, 69% of the total population forms the state's agricultural workforce (Govt. of Assam, 2003). However, in the past few decades, there has been a rehabilitation of communities along with rapid urbanization and industrial growth which has taken a toll on the surface water quality of the state. According to a report published by United Nations Development Program (UNDP) in 2004, unplanned urbanization and population growth not only had several implications to food security and crop yield, there has been a continuous depletion and degradation of the surface water quality in the state. Furthermore, elevated concentrations of arsenic and fluoride in the groundwater increases the dependency of its population on surface water sources for drinking, household, and irrigation purposes. Therefore, the state of surface water sources must be such that they meet water quality standards for drinking. This calls for strict water quality monitoring programs, which would thereby help in understanding the status of water bodies. This would aid in evolving policies to diminish the additional stresses on surface water resources.

1.4 SCOPE OF THE STUDY

Scope of the study is confined to the reconnaissance survey of study area, identification of sampling sites, collection of surface water samples from surface water resources (Brahmaputra River and its seven tributaries, and Deepor Beel) and assessment of observed water quality data sets. Water samples were analyzed for various physico-chemical parameters including heavy metals. CA was applied on the observed data sets to determine the site similarities. DA was employed to identify the discriminating parameters and finally PCA was applied to determine the possible sources of pollution. Entropy weighted WQI was used to determine the spatial variation of surface water quality. Furthermore, based on information entropy heavy metal contamination index and irrigation water quality index have been proposed. Risk assessment due to heavy metal contamination was evaluated using Hazard Quotient (HQ), developed by United States Environment Protection Agency (USEPA). Finally, a novel technique has been proposed for the identification of ideal sampling sites using Diversity index and Disorder index.

1.5 THESIS ORGANISATION

The thesis is organised chapter wise as below:

- Chapter 1 deals with the surface water quality monitoring and assessment. Objectives, need of the present study, and scope of the study are also discussed in this chapter.
- Chapter 2 provides a detailed literature review on the various factors affecting surface water quality, application of CA, DA and PCA for source identification of pollutant, various methodologies adopted for surface water quality evaluation, and application of information entropy in surface water quality monitoring and assessment.
- Chapter 3 is associated with the approach of the various phases of the research, study area and sampling strategy. Various methodologies adopted for present study are also discussed in this chapter.
- Chapter 4 provides a discussion on the observed water quality data-sets.
- Chapter 5 deals with the application of various MSTs and source identification of pollutants.
- Chapter 6 confers the application of information entropy in surface water quality assessment.
- Chapter 7 deals with the identification of ideal monitoring sites.
- Chapter 8 concludes the present study and provides recommendations for future work.

2

Chapter

Literature Review

This chapter covers review on the

- Water quality monitoring
- Multivariate statistical techniques
- Water Quality Indices
- Information entropy
- Application of information entropy in water quality and hydrological studies

---

2.1 BACKGROUND

Water is one of the most important substances on earth. Clean and safe water (Fig. 2.1) is necessary for the survival of life. Only 2.5% of water on the Earth is fresh water, and over two-thirds of this is frozen in glaciers and polar ice caps. 87% of the remaining fresh water is contained in lakes, 11% in swamps, and only 2% in rivers (Fig. 2.2). Surface water sources are of immense importance geologically, biologically, historically and culturally. It acts as a source of water for various domestic purposes, provides travel routes for exploration, commerce and recreation, host biologically diverse ecosystems provide staple diet and occupation to local people. Due to technological and industrial progress, nature of the aquatic environment in urban areas undergoes numerous changes and deteriorating its quality. In the endeavor to manage water to meet human needs, the needs of freshwater species and ecosystems are neglected.

In a developing country such as India, point source of pollution such as the discharge of domestic and industrial wastewater, landfill leachate etc. and non-point sources such as agriculture and urban runoff pose a major challenge for pollution control boards and decision makers. The discharge of oxygen-demanding substances, toxic wastes, suspended solids and artificially coloured wastes affect both the quality and the aesthetics of water resources. Eutrophication caused by excessive discharge of nutrients (Nitrogen and

Phosphorous) impairs a waterbody from uses such as recreation, drinking, irrigation and fishing (Carpenter et al., 1998). The emergence of trace contaminants and microplastics in the environment pose a serious threat to aquatic environments and human life (Wang et al., 2017). The use of pesticides at a global level due to intensive agriculture coincides with the accumulation of pesticide residues in aquatic ecosystems. Pesticide residues reach aquatic bodies through run-off, leaching, and thoughtless disposal of empty pesticide containers (Konstantinou et al., 2006). The natural balance in the aquatic ecosystem, which must be maintained, gets disturbed due to various natural and anthropogenic activities.

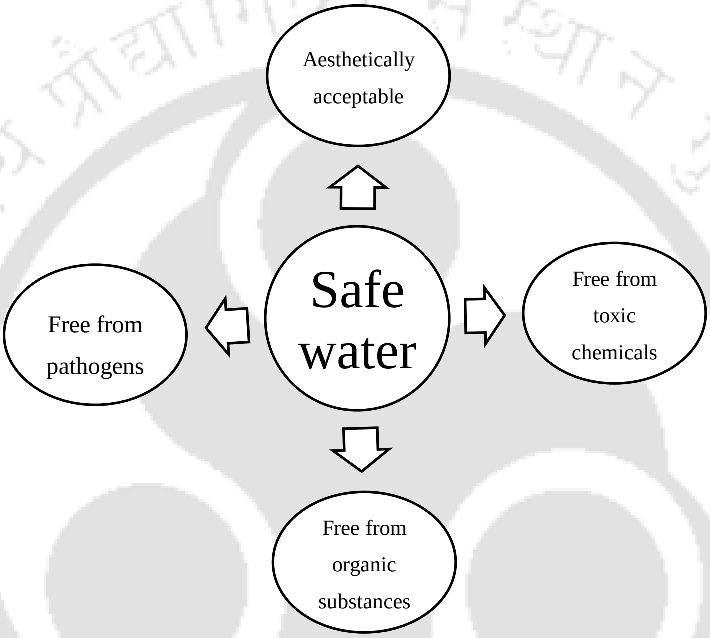


Fig. 2.1. Characteristics of safe water

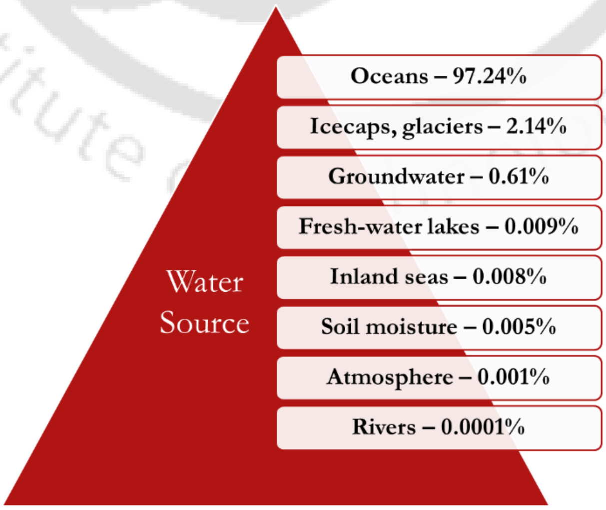


Fig. 2.2. Distribution of water on earth (Source: USGS, [1984]- The Hydrologic Cycle (Pamphlet))

2.2 WATER QUALITY MONITORING

Water quality monitoring programs are essential to regulate and control a wide network of water resources. Water quality monitoring may be defined as any effort undertaken by government or private enterprise in understanding the physical, chemical and biological characteristics of water by statistical sampling. Monitoring is performed for effective water resources planning, management and development of decision making strategies for preservation and stream restoration (Ward et al., 1986).

The main objective of the design of monitoring program is to reduce the budget without compromising the anticipated information to the level of precision and accuracy. The scoping and designing step is a key element of the entire water quality monitoring program, which is based on clear scientific understanding of: 1. problems; 2. relevant background information; 3. monitoring objectives; 4. desired outcomes; 5. suitable methods; 6. the dynamics and characteristics of water systems.

Water quality monitoring involves eight crucial steps (Fig. 2.3). First step is setting water quality monitoring objectives. Objectives of water quality monitoring may vary depending on the specific task or project (Zhang, 2007). Objectives may be:

(a) To determine the pollution level of water body or concentration of one or more pollutant
(b) To identify the latent pollution sources
(c) To measure the ambient background concentration.
(d) To study the fate and transport of contaminants
(e) To detect the accidental release
(f) To determine the changes in water quality
(g) Scientific research

Second step is the assessment of the availability of resources. Resources for water quality monitoring include laboratory facility, transportation facility and adequate number of skilled manpower. Third step is reconnaissance survey of monitoring area. The purpose of survey is to observe all sources of wastes, entering tributaries that might contribute a potential pollutant threat. Survey also encompassed the collection of background information such as hydrology, land use, urbanization, industrialization and agriculture etc. in catchment area. Fourth step in water quality monitoring is design of network. Design of network includes the selection of sampling locations, frequency of sampling, parameters to be measured. Fifth step involves the collection of water samples. For water sampling cleaned bottles, free from dust and dirt is required. Sample should be representative of the overall water quality. It is essential to keep the sample free from any contamination. Next step is the analysis of samples in laboratory. For the analysis of water samples standard methods should be followed throughout the analysis. Deionized water is required for carrying out the dilutions. A quality control procedure must be maintained throughout analysis, including recalibration of instruments. Last step in surface water quality monitoring is quality assurance (QA). Aim of QA is to acquire an estimate of the precision and accuracy of the data.

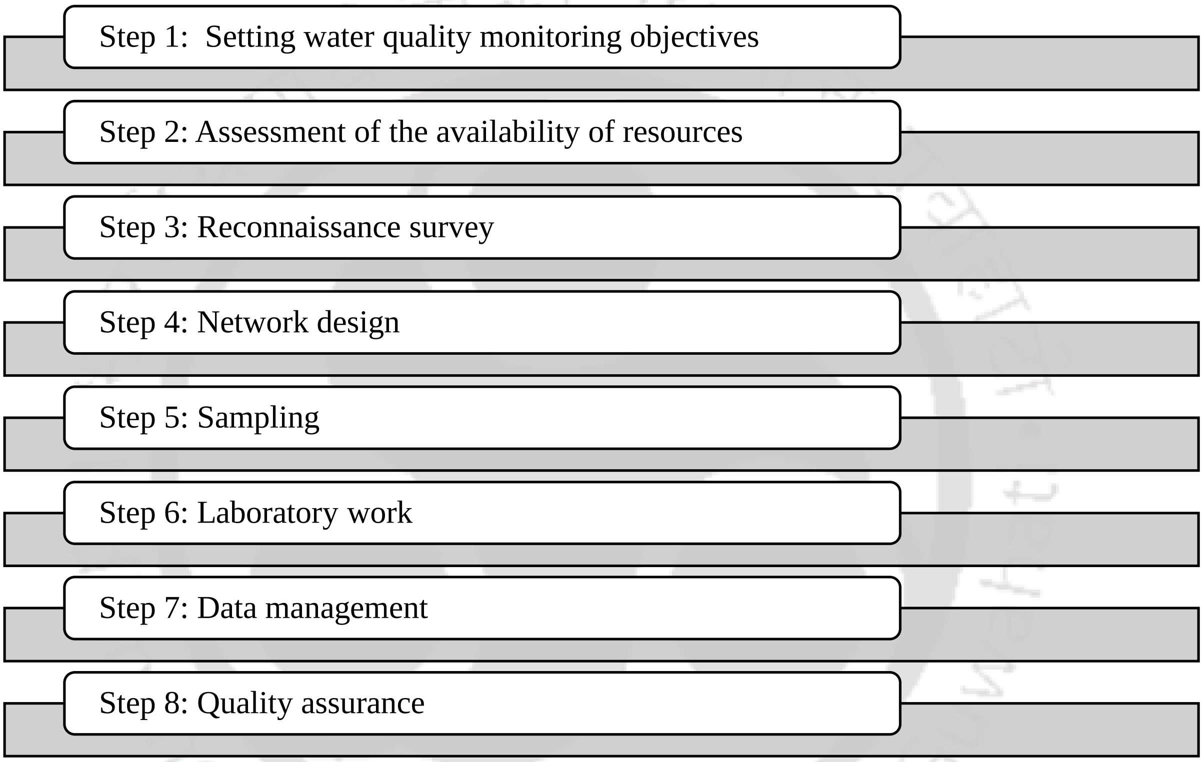


Fig. 2.3. Steps involved in water quality monitoring program

2.3 MULTIVARIATE STATISTICAL TECHNIQUES (MSTs)

The application of different multivariate statistical techniques (MSTs), such as cluster analysis (CA), discriminant analysis (DA) principal component analysis (PCA) and factor analysis (FA), facilitates the interpretation of complex datasets to better understand the water system. It also allows the identification of possible factors/sources that influence water quality and offers a valuable tool for the reliable management of water resources as well as rapid solutions to pollution problems (Vega et al., 1998; Simeonova et al., 2003; Shrestha and Kazama, 2007; Varol and Şen, 2009).

2.3.1 Cluster Analysis

Cluster analysis (CA) is a multivariate technique and effective tool for determining the structure within complex datasets and for exploring similarities and/or dissimilarities among data and variables (Singh et al., 2004; Tsai et al., 2016). CA divides the data into groups or clusters that are meaningful and useful. Primary purpose of CA is to assemble objects based on the their characteristics (Shrestha and Kazama, 2007). Various methods for clustering can be divided in four groups.

(a)	Partitional clustering:	Directly divides data points into some pre-specified number of clusters without a hierarchical structure.
(b)	Hierarchical clustering:	Groups data with a sequence of nested partitions, either from singleton clusters to a cluster containing all elements.
(c)	Density based clustering:	Divides the feature space in order to find dense areas separated by empty space.
(d)	Grid based clustering:	Based on a multiple-level granularity structure.

The merge of a pair of clusters or the formation of a new cluster is dependent on the definition of the distance function between two clusters. The proximity of two groups of elements is defined as linkage. There are three methods of linkages (Fig. 2.4):

(a) Single Linkage (Nearest neighbor method)

(b) Complete Linkage (Farthest-neighbor or diameter method)

(c) Average Linkage (Centroid method)

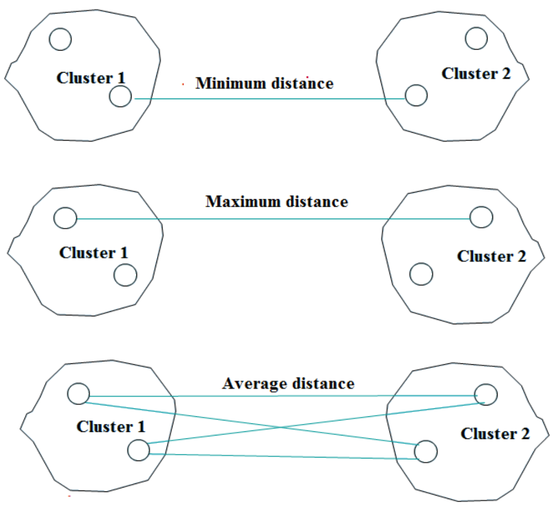


Fig. 2.4. Linkage of cluster

Hierarchical agglomerative clustering (Ward's method of linkage (Fig. 2.5) with Euclidean distance as a measure of similarity) is the most common approach, and is typically demonstrated by a dendrogram (Fig. 2.6), which provides a visual summary of the clustering processes (Einax et al., 1998 and McKenna, 2003). Ward's method joins the two clusters A and B that minimise the increase in the sum of squared errors (SSE). SSE within and between clusters (cluster A and cluster B) can be calculated as:

$$SSE_A = \sum_{i=1}^{n_A} (a_i - \bar{a})' (a_i - \bar{a}) \tag{2.1 a}$$

$$SSE_B = \sum_{i=1}^{n_B} (b_i - \bar{b})' (b_i - \bar{b}) \tag{2.1 b}$$

$$SSE_{AB} = \sum_{i=1}^{n_{AB}} (y_i - \overline{y_{AB}})' (y_1 - \bar{y}_{AB}) \tag{2.1 c}$$

where a_i represents the i^{th} observation vector in cluster *A*, and $\bar{a}$ is the centroid of cluster *A*, b_i represents the i^{th} observation vector in cluster *B*, and $\bar{b}$ *is* the centroid of cluster *B* and y_i represents the i^{th} observation vector in cluster *AB*, and $\bar{y}_{AB}$ the centroid of newly formed cluster *AB*.

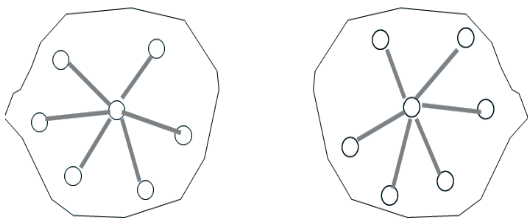


Fig. 2.5. Ward's Linkage of cluster

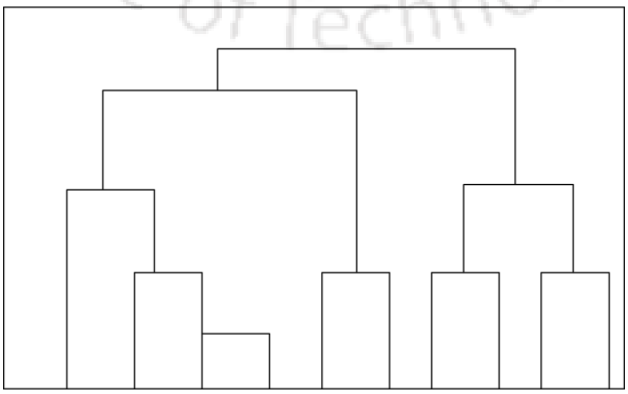


Fig. 2.6. Dendrogram

2.3.2 Discriminant Analysis

Discriminant analysis (DA) is a statistical analysis which is used to determine which continuous variables discriminate between two or more naturally occurring groups. The mathematics of DA are very similar to the one-way MANOVA. In fact, the roles of the variables are simply reversed. The classification variable in the MANOVA becomes the dependent variable in DA. The dependent variables in the MANOVA become the independent variables in the DA.

DA is a seven step process. First step is collection of training data. Training data are data with known group memberships. Second step is calculation of prior probabilities. The prior probability represents the expected portion of the community that belongs to population. Third step involves the use of Bartlett's test to determine if the variance-covariance matrices are homogeneous for all populations involved. Fourth step is computation of discriminant functions (DF). DF can be expressed as (Shrestha & Kazama, 2007):

$$f(G_i) = k_i + \sum_{j=1}^{n} w_{ij} P_{ij} \tag{2.2}$$

where *i* is the number of groups *(G)*, k_i is the constant inherent to each group, *n* is the number of objects used to classify a set of data into a given group, w_{ij} is the weight coefficient, assigned by DA to a given selected parameter (P_j). This function is used to classify the new object into one of the known population. Next step is the use of cross validation to estimate misclassification probabilities. Last step is classification of observations with unknown group memberships.

2.3.3 Principal Component Analysis

PCA is a statistical method of data reduction (Yim et al., 2015). It is based on the correlation matrix in which the intercorrelations among the variables are presented. Main aim of PCA is to bring intercorrelated variables under more general, underlying variables. PCA is also a variable reduction techniques and is used when variables are highly correlated. It is far more commonly used than Factor Analysis (FA) (Bouguerne et al., 2017). It is a way of classifying patterns in data, and expressing the data in such a way as to highlight their similarities and differences. The principal component (PC) can be expressed as (Shrestha and Kazama, 2007):

$$z_{ij} = a_{i1}x_{1j} + a_{i2}x_{2j} + \cdots \ldots \ldots \ldots \ldots \ldots . a_{im}x_{mj} \tag{2.3}$$

where z is the component score, a is the component loading, x the measured value of variable, i is the component number, j the sample number and m the total number of variables.

2.4 STUDIES ON MSTs USED FOR WATER QUALITY ASSESSMENT

Simeonov et al. (2003) applied MSTs for the assessment of surface waters in Northern Greece. The data set consisted of analytical results from a 3-yr survey conducted in the major river systems (Aliakmon, Axios, Gallikos, Loudias and Strymon) as well as streams, tributaries and ditches. CA was applied to identify similarity groups between the sampling sites. In this study hierarchical agglomerative cluster analysis was performed on the normalized data (z score) using Euclidean distance as similarity measure with Ward's method of linkage. The significance of the clusters obtained was tested by the Sneath's index of disjunction. CA grouped all the sampling sites in four cluster. PCA was performed on data set to compare the compositional patterns between the examined water systems and to identify the factors that influence each one. PCA yielded six PCs with eigenvalues >1 summing almost 90% of the total variance.

Singh et al. (2004) evaluated the of temporal and spatial variations in water quality of Gomti River in Northern part of India using MSTs. Gomti River was monitored at eight different sites regularly over a period of 5 years (1994–1998) for 24 parameters. Obtained data sets were treated with CA, DA and PCA. CA classified the eight monitoring sites into three statistically significant clusters. PCA of the entire data set resulted six PCs with eigenvalues >1 explaining about 71% of the total variance in the water-quality data set. DA used only five parameters to discriminate between the seasons with 88% correct assignations and nine parameters to discriminate between the three spatial regions with 91% correct assignations.

Shrestha and Kazama (2007) applied CA, DA and PCA for the assessment of water quality data set of the Fuji river basin. In this study hierarchical agglomerative cluster analysis (Ward's method of linkage and squared Euclidean distances as a measure of similarity) was performed on the normalized data set. HCA grouped 13 sampling sites into three clusters, based on the similarity of water quality characteristics. DA was performed for both spatial and temporal analysis. PCA was applied to the data sets of the three different groups resulted from CA. PCA resulted in five, five and three latent factors explaining 73.18, 77.61 and 65.39% of the total variance in water quality data sets of all three clusters.

Boyacioglu and Boyacioglu (2007) investigated the surface water quality of Buyuk Menderes River and its tributaries using CA and Factor analysis (FA). For evaluation of surface water quality eleven variables were measured at 17 sampling sites. FA resulted three factors explaining 80.6% of total variance. CA grouped all the sampling locations in three cluster. Cluster 1 included sampling locations from the downstream of the Buyuk Menderes River Cluster 2 consisted sites that are in the middle and at the east part of the basin where agriculture is heavily practiced and Cluster 3 included three stations, two out of which situated at the downstream and one in the middle part of the basin.

Yidana et al. (2008) applied R-mode hierarchical CA and PCA to surface water quality data from Ankwaso, Dominase and Prestea, along the Ankobra Basin, Ghana. The result of the study indicated that hydrochemistry at Ankwaso was mainly affected by the weathering of silicate minerals, whereas those of Dominase and Prestea were, respectively, influenced by precipitation, domestic wastewaters, and the decay of organic matter.

Kazi et al. (2009) evaluated the water quality of Manchar Lake (Pakistan), using CA and PCA. Lake was monitored at five different sites for 36 parameters during 2005–06. CA classified five sampling sites into three clusters. Three PCs resulted from PCA analysis explained 97.6% of the variance in the data set of lake water. Study concluded that the major sources of water quality deterioration were the inflow of effluent from industrial, domestic, agricultural and saline seeps into the lake and from people living in boats and fishing.

Krishna et al. (2009) applied MSTs for the evaluation of heavy metals in surface and groundwater of Patancheru industrial town near Hyderabad (India). Data sets included thirteen parameters from 53 sampling sites. R-mode FA and PCA was used for the data treatment. FA resulted four factors responsible for data structure explaining 75% of total variance in surface water and two factors in groundwater explaining 85%. This study indicated the usefulness of MSTs for evaluation and interpretation of the hazardous toxic elements data to get better information.

Varol and Sen (2009) applied MSTs for the assessment of water quality data set of the Behrimaz Stream (Turkey). In this study, hierarchical agglomerative CA was performed on the normalized data set by means of the Ward's method, using Euclidean distances as a measure of similarity. Hierarchical CA grouped four monitoring sites into two groups, relatively less polluted (LP) and medium polluted (MP) sites. FA of the two data sets resulted five factors for the both clusters with Eigenvalues >1, explaining 88.32% and 88.93% of the total variance. Based on analysis, study concuded that the parameters responsible for water quality variations are mainly related to discharge, temperature, and soluble minerals and nutrients in LP sites; and organic pollution and nutrients in MP sites in the basin.

Bu et al. (2010) applied CA, DA, and FA to evaluate water quality of the Jinshui River of China. CA grouped 12 sampling sites with 22 variables into three clusters reflecting the different pollution levels. DA confirmed the three clusters with nine discriminant variables. FA extracted five PCs explaining 90.01% of the total variance and representing chemical component, oxide-related process, natural weathering and decomposition processes, nutrient process, and physical processes, respectively.

Varol et al. (2012) applied CA, PCA, FA and DA to assess the temporal and spatial variations of water quality data sets for Kralkızı, Dicle and Batman dam reservoirs (Tigris River basin) obtained during 1 year (2008–2009) of monitoring. Hierarchical CA classified 12 months into two clusters and grouped ten monitoring sites into four clusters based on similarities in the water quality characteristics. PCA/FA extracted five PCs in the data structure explaining 80% of the total variance. Temporal DA revealed nine parameters giving 100% correct assignations. Spatial DA revealed eight parameters giving 92.5% correct assignations.

Barakat et al. (2016) assessed the spatial and seasonal water quality variation of the Oum Er Rbia River and its main tributary, El Abid River and determined the main contamination sources using MSTs. The water quality data were collected during 2000–2012 from fourteen sampling sites distributed along the river. The water quality parameters used were temperature, pH, EC, turbidity, TSS, DO, NH_4^+, NH_3^-, TP, BOD_5, COD and FC. PCA extracted four PCs explaining 63% of the total variation in the data. Result indicated that the variations in water quality were mainly related to point source contamination (domestic and industrial wastewater), non-point source contamination (agriculture activities), as well as natural processes (weathering of soil and rock).

Hajigholizadeh and Melesse (2017) assessed the spatial and temporal variations of water quality of the Miami Canal, Kissimmee River and Caloosahatchee River, in South Florida using CA and DA. 15 years (2000–2014) data sets of 12 water quality parameters were collected from of 16 monitoring stations. Agglomerative hierarchical CA grouped the all monitoring sites into three clusters (low pollution, moderate pollution, and high pollution). Stepwise DA identified chl-a, DO, TKN, TP and temperature as the most important discriminating water quality parameters responsible for temporal variations. For spatial variation in wet season chl-a, DO, TKN, TP, Mg^{2+}, Cl^-, and Na^+ were the most important discriminating parameters and in dry season, DO, TKN, TP, turbidity, Mg^{2+} and Cl^- were most important discriminating parameters.

Ling et al. (2017) applied MSTs for the assessment of surface river water quality of Batang Baramand river and its tributaries in Sarawak (Malaysia). Twenty water quality parameters were measured at thirty locations in the year 2015. CA grouped all the sampling sites into four clusters representing the upstream, middle, and downstream regions of the main river and the tributaries from the middle to downstream regions of the river. PCA resulted six PCs that explained 83.6% of the data set variance. The first PC was associated with the logging activities. Second PC was associated with the discharges from domestic wastewater.

Luo et al. (2017) assessed the effects of urbanization on the water quality of 19 rivers in Liangjiang New Area, China, using MSTs. Data collected in April (dry season) and September (wet season) of 2014 and 2015. CA grouped 19 sampling sites into two clusters, which located at sub-catchments with high- and low-level urbanization, respectively. FA identified the five factors of water quality parameters, which explain majority of the experimental data. Nutritious pollution, seasonal changes, and construction activities were three major factors affecting the water quality of monitored rivers.

Wang et al. (2017) assessed the origins of the dissolved trace elements in the Huaihe River (Anhui, China) with the application of different MSTs, including correlation matrix CA and FA/ PCA. Water samples were collected from 53 sampling locations from 5–10th July 2013 at four different depths (0, 2, 4 and 8 m). CA was applied to classify the all sample sites into three groups of water pollution, high pollution, moderate pollution, and low pollution, reflecting influences from tributaries, power plants and vehicle exhaust, and agricultural activities, respectively. FA/PCA identified three source types that accounted for 79.31% of the total variance.

Zhang et al. (2018) assessed temporal and spatial variations in water quality of the Jialing River watershed in Guangyuan City, China using MSTs. Temporal CA grouped the months into two clusters with similar physiochemical water quality characteristics. First cluster included May–November months, (high-flow and mean-flow periods). 2nd cluster grouped the months of the low-flow period. DA identified, T, discharge (Q), DO, oils, F, and Cd as the discriminant variables. The spatial similarity analysis showed that the sampling stations can be divided into two clusters as highly polluted regions and low polluted regions

2.5 WATER QUALITY INDICES (WQIs)

Water quality index (WQI) is an important tool to summarize the water quality data to a numerical score for easy understanding of users and policy makers (Abbasi and Abbasi, 2012; Abtahi et al., 2015; Bouguerne et al., 2017). It expresses the water quality in the form of index number. First modern WQI was developed by Horton (1965). Since then various researchers have developed and applied different water quality and pollution indices to classify the water quality in their region (Abtahi et al., 2015; Bouguerne et al., 2017). Steps involved in development of any WQI is shown in Fig. 2.7 (Abbasi and Abbasi, 2012).

Step 1

- Selection of water quality parameters

Step 2

- Transformation of selected parameters to a common scale

Step 3

- Assignment of weightages to all the selected parameters

Step 4

- Aggregation of subindices

Fig. 2.7. Development of WQI

Selection of water quality parameters is a critical step in the development of WQI. It is essential to select a few important parameters among a significant number, representative of the water quality for particular use. In most of the WQIs parameters is decided through either expert judgement or through the statistical methods (Sutadian et al., 2015). Different water quality parameters are expressed in different units and different ranges. Before the formulation of index, all the water quality parameters have to transferred into a common scale (sub-indexing). After the transformation of parameters into a single scale, weight is assign to all selected parameters with regard to their relative importance. Finally, sub-indices are aggregated to develop the final index.

2.6 STUDIES ON DEVELOPMENT OF WQIs

Horton's Index: Horton (1965) proposed a WQI based on eight parameters (sewage treatment, DO, pH, Coliform density, specific conductance, carbon chloroform extract (CCE), alkalinity and chlorides). Rating scales 0 to 100 were assigned to each parameters. Parameters were weighed (weighting factor 1 to 4) according to its relative importance. The resultant water quality index had values in the range from 0 to 100 with higher values signifying a better quality and vice versa. Horton's index is easy to compute but highly subjective (Abbasi and Abbasi, 2012).

NSF – WQI: This index was developed by Brown et al. (1970) and supported by National Sanitation Foundation (NSF). NSF-WQI was based on Delphi technique. For formulating NSF-WQI, Brown et al. formed a panel of 142 persons from various professions throughout the United States with expertise in various aspects of water quality management. Based on the experts' opinion, the researchers selected nine individual parameters and two grouped parameters of greatest significance. The nine individual parameters were DO, Fecal Coliform (FC), pH, BOD_5, NO_3^-, PO_4^{3-}, Temperature, Turbidity, and Total Solids (TS). The two grouped variables selected were Toxic Substances and Pesticides. Weight of the parameters were also assigned by the opinion of experts. the final weights (in brackets) were as follows: DO (0.17), FC (0.16), pH (0.11), BOD_5 (0.11), temperature (0.10), PO_4^{3-} (0.10), NO_3^- (0.10), turbidity (0.08) and TS (0.07). NSF-WQI is general water quality index and does not incorporate any specific water function.

Prati's Implicit Index of Pollution: This index was proposed by Prati et al. (1971). In the first step, water quality was classified vis a vis all the parameters based on quality standard. In the second step, one pollutant was taken as reference and its actual value was considered directly as reference index. In next step mathematical expressions were developed to transform each of the values of the other pollutants into subindices. Finally, index was calculated as the airthmatic mean of all subindices.

Deininger and Landwehr (1971) developed a WQI for public water supply. In this index 11 water quality parameters were employed for surface water and 13 parameters were employed for groundwater. Two aggregation function, additive and geometric mean were used for surface and groundwater respectively. Weight of all parameters were slightly different from NSF-WQI.

Oregon WQI (OWQI): OWQI was developed by Oregon Department of Environmental Quality (ODEQ), USA. Delphi technique was employed for the selection of water quality parameters and it was based on water quality data of the Willamette River basin in Oregon. Eight selected parameters (temperature, DO, BOD, pH, ammonia + nitrate nitrogen, total phosphorus, TS, and FC) were integrated to develop a final index. OWQI was updated by Cude (2001) by refining the original sub-indices and improving the aggregation method.

British Columbia WQI (BCWQI): BCWQI was based on a variety of objectives depending on the designated water use: drinking, recreation, irrigation, livestock watering, wildlife and aquatic life. Separate rankings were published based for each use. BCWQI was based on three factors. First factor represents the percentage of water quality guidelines exceeded. Second factor was percentage of measurement in which one or more of the guidelines were exceeded. Third factor was the maximum extent by which any of the guidelines were exceeded.

Canadian Council of Ministers of Environment Water Quality Index (CCME-WQI): CCME-WQI adopted the conceptual model from British Columbia index (BCWQI). The index is based on a combination of three factors:

(a) Scope (F1): It represents the extent of water quality guideline non-compliance over the time period of interest. (Adopted from BCWQI).

(b) Frequency (F2): It represents the percentage of individual tests that do not meet objectives ("failed tests"). (Adopted from BCWQI).

(c) Amplitude (F3): It represents the amount by which failed test values do not meet their objectives.

Stochastic Water Quality Index (SWQI): Lohani and Mustapha (1982) proposed a SWQI incorporating the probabilistic water quality entities. This index was based on assumption that water quality parameters are random variable, hence it can be deal better with probability theory. SWQI was applied to the Linggi River in Malaysia for classification of its water quality.

Smith's Index: Smith (1990) proposed a new index based on expert opinion and water quality standards. This index addressed four type of water application i.e. general, public bathing, water supply and fish swamping.

Overall Index of Pollution (OIP): Based on a general classification scheme in Indian context, OIP was developed by Sargaonkar and Deshpande (2003). Water quality parameters considered in the present study were pH, Turbidity, Color, DO, BOD_5, Cl^-, TDS, Hardness, SO_4^{2-}, NO_3^-, As, F^- and TC. OIP was estimated using following mathematical expression:

$$OIP = \frac{\sum P_i}{n} \tag{2.4}$$

where Pi is the pollution index of i^{th} parameter and n is the total number of parameters. Pi of all water quality parameters was calculated using the mathematical equations given by Sargaonkar and Deshpande (2002).

Global Stochastic Quality Index (GSWQI): Beamonte et al. (2005) proposed GSWQI considering the uncertainty surrounding the quality classification still remaining after the data had been observed. This index was built with the probability classification vector of each water quality parameter. In order to obtain probability classification vector, a mixed-lognormal model was introduced. GSWQI was applied to the data observed in the La Presa sampling station, located on the river Turia, near Valencia city (Spain).

Hybrid Fuzzy-Probability WQI: Nikoo et al. (2010) developed a hybrid probabilistic WQI by incorporating fuzzy interface system, Bayesian network and probabilistic neural network. Index was applied on the water quality data of Jajrood River (Iran). Result proved its effectiveness in assessment and zoning of river water quality.

Dynamic WQI (D-WQI): Feng et al. (2015) presented a dynamic WQI (D-WQI) model based on functional data analysis. D-WQI was a generalization of the conventional WQI. In the D-WQI model, the changes of water quality and pollutant importance were characterized in the form of dynamic functional curves. Further, D-WQI model was applied to the water quality assessment of the Changjiang River in Sanjiangying.

Lobato et al. (2015) proposed a new Quality Indicator and WQI for the assessment of the water quality of a Tucuruí Hydroelectric Plant reservoir in the Amazon area (Brazil). FA was used to select the appropriate water quality parameters. This index was based on NSF-WQI.

Whittaker's Index: Whittaker et al. (2015) proposed a WQI based on application of index number theory. The relative weights of the index were articulated endogenously, from the data itself.

WQImin: Sun et al. (2016) proposed a modified WQI (WQImin) based on PCA and correlations analyses of the water quality parameters. WQImin was used as an indicator to assess the spatial and temporal trends of water quality in the Dongjiang River in southern China.

Bottled Water Quality Index (BWQI): BWQI was developed by Tsakiris et al. (2017). The index combined one microbiological and six physicochemical parameters. The assessment was based on both laboratory analysis results and values on bottle labels. This index was developed in two steps. The first step was a ON/OFF process for testing each bottled water sample against the *E. coli* concentration and the permissible range of pH. The second step was a multiplicative model, which includes three types of the most representative quality parameters for bottled water assessment. The BWQI is a quick, sensitive and powerful tool to assess and compare the quality of bottled waters.

Functional water quality index (FWQI): Sahoo and Patra (2018) developed FWQI for the evaluation and classification of surface water quality at River Basin system. Fuzzy entropy weight method was used to calculate the weight of selected parameters. The assessment of FWQI was based on the presumption that each indicator may have separate as well as cumulative potential to degrade the quality of water which causes an adverse effect to ecological and environmental health.

West Java Water Quality Index (WJWQI): WJWQI was proposed by Sutadian et al. (2018) to evaluate water quality in rivers of the West Java Province, Indonesia. In this index, selection of parameters was based on a statistical assessment. Sub-index functions were developed using the permissible limits for various levels of intended uses or water classes. APH method was used to determine the weight of selected parameters. Finally, non-equal geometric method was used to develop the final index

Neuro-fuzzy based WQI: Yaseen et al. (2018) proposed a hybrid adaptive neuro-fuzzy models for estimation of WQI. Different hybrid intelligence models based on adaptive neuro-fuzzy inference system integrated with fuzzy c-means data clustering, grid partition and subtractive clustering models were used in modelling river WQI. In this study, the ability of three different neuro fuzzy methods in modeling water quality parameters were examined for the Selangor river basin.

Application of WQI to evaluate the water quality status has been adopted by many organizations and agencies, but there is no globally accepted models for developing a WQI. Based on literature, it was observed that all WQIs have their own advantages and disadvantages. Some studies are available on the review on the existing WQIs. Fernandez et al. (2004) compared thirty-six WQIs on the same sample of water. Lumb et al. (2011) reviewed the genesis and development of various WQI models developed from the 1960s till 2010 and presented the importance of various WQIs, the steps used in their formulation and their current uses. Tyagi et al. (2013) reviewed four popular WQIs and explained their merits and demerits. Sutadian et al. (2016) reviewed thirty existing WQIs and identified seven as most important based on their broader application. However, there is no systematic and thorough review of existing WQIs in the literature to explore and assess the steps used in their development

2.7 INDICES FOR HEAVY METALS AND THEIR APPLICATION

Due to toxicity, persistence and bioaccumulation, quantification of heavy metal contamination and assessing the suitability of water for drinking purposes calls for proper indexing of heavy metal concentrations and aggregating them into a numerical score (Pekey et al., 2004). Heavy metal pollution indices have been used by various researchers to evaluate the heavy metal contamination in ground water and surface water (Mohan et al., 1996; Backman et al., 1998; Prasad et al., 2014; Tiwari et al., 2015; Abdel-Satar et al., 2017; Chaturvedi et al., 2018). The most popular indices are Heavy Metal Pollution Index (HPI), Contamination Index (CI) and Heavy Metal Evaluation Index (HEI). Such indices are dimensionless number and used for classification of water resources in term of contamination. Although, these indices are convenient for quantification of heavy metal contamination and spatial-temporal assessments yet they contradict each other in some circumstances (Bhuiyan et al., 2010; Singh et al., 2017).

The heavy metal pollution index (HPI) was developed by Mohan et al., (1996). HPI has been evaluated using the weighted arithmetic average method of indexing. Mohan et al. (1996) applied the HPI on four heavy metals (Cu, Pb, Zn and Cd) at twelve residential areas. Prasad and Bose (2001) evaluated the HPI at eight surface water locations and nine post-monsoon water locations near a limestone mining area in Sirmour district of Himachal Pradesh. The HPI was evaluated against the concentrations of seven heavy metals namely Cu, Cd, Cr, Fe, Mn, Pb and Zn. The critical index for drinking water was suggested as 100. The HPI at all the locations were found to be well below the critical index limit of 100 suggesting the sources were not contaminated with respect to heavy metals. Prasad and Sangita (2008) evaluated the HPI of groundwater samples of an open cast mine filled with fly ash. Although the average Mn concentration (1.706 mg/L) in this study was much higher than the highest permissible standard (0.3 mg/L), the maximum HPI (75.10) was found well below the critical limit of 100. In another study, Giri et al., (2010) evaluated the HPI with respect to the heavy metal concentrations at ten ground water locations and eight surface water locations in a mining and mineralised zone of East Singhbhum region (India). The HPI of both groundwater and surface water came to be below 100. Kumar et al., (2012) assessed the heavy metal contamination in ground water of Chennai, India using HPI in conjunction with spatial distribution maps. The highest HPI was value was 97.55 which was below the critical index limit (100) given by Prasad and Bose (2001). The classification scale in the study was modified in accord with the HEI classification scale proposed by Edet and Offiong (2002). The new gradation scale proposed nine samples to be low contaminated (HPI < 19), three samples as medium contaminated (HPI- 19-38) and two samples to be highly contaminated (HPI > 38). In a study conducted by Giri and Singh (2014), 21 surface water samples from Subarnarekha river during pre-monsoon and post-monsoon periods were assessed for heavy metal pollution with respect to 12 heavy metals. The HPI of the river varied highly from 3.55 to 388.9 due to the significant seasonality of the heavy metal concentrations. The HPI was also reclassified as per the scale put forward by Edet and Offiong (2002). Tiwari et al. (2015) evaluated the water quality of 28 water samples from 14 sites of West Bokaro coalfield, India using GIS techniques and HPI. Although, the HPIs of the water samples were below the critical index limit yet the reclassification done suggested HPI < 15 as less contaminated and HPI between 15 and 30 as medium contaminated. Qu et al., (2018) evaluated the heavy metal pollution and potential ecological risks to human health from the concentration data of seven heavy metals from 14 locations spanning the rural-urban interface of the Wen-Rui Tang river watershed in southeast China. All the HPI values were below 100 and a transition was noted in the HPI which exceed 80 till 2005 to below 40 after 2005.

The evaluation of the CI was developed at the Geological Survey of Slovak Republic by Bodis and Rapant and was further refined at the Geological Survey of Finland (Backman et al., 1998). Individual components or parameters exceeding their upper permissible limits were aggregated for the calculation of CI for a sampling site. The combined effects of the toxic metals were thus summarized to a single numeric value. In another study conducted by Edet and Offiong (2002), CI and HPI were applied at Akpabuyo-Odukpani area, Lower Cross River Basin (southeastern Nigeria). The two indices gave extreme classification results although they depicted good correlation with the heavy metal concentrations. Subsequently, a heavy metal evaluation index (HEI) was proposed. The HEI developed to overcome deficiency in the conflicts between HPI and CI, classified water quality with respect to multiple of the mean as criteria. Furthermore, the HPI and CI scales were modified with respect to their multiples of mean. Such modifications deemed it possible to compare the three indices, however, the HEI was adopted due to its simplicity. In a study conducted by Bhuiyan et al., (2010), hazardous materials evaluation had been done in irrigation and drinking water systems in the vicinity of a coal mine area in northwestern Bangladesh using HPI and CI. HPI and CI yielded different results despite significant correlations between them. The HEI showed strong correlations with HPI and CI. Modifications to the existing HPI and CI values depicted comparable results with HEI. Prasanna et al. (2012) conducted study at Curtin Lake, Miri City, East Malaysia for assessing the heavy metal pollution using HPI, CI and HEI and concluded that the HEI depicted strong correlations with HPI and CI. The existing gradation scales of the indices were also modified using a multiple of mean. Singh et al., (2017) assessed the groundwater contamination of resources by potentially toxic trace elements near a coal mining area of the Korba coalfield, Central India. A unified approach of pollution evaluation indices (HPI, HEI and CI) and statistical techniques were applied to the results. The indices contradicted each other and the results were in ambiguity in order to assess the heavy metal pollution levels in the groundwater regime of the active mining area. Consequently, a multiple of the mean values approach had been adopted to remove the conflicts between the indices and reclassify them according to this scale.

Recently, a modified HPI developed by Chaturvedi et al., (2018) tried to overcome the shortcomings of the HPI by introducing a modified sub-indexing approach of removing the modulus from the sub-index (Q_i) and introducing a positive index (PI) and a negative index (NI) for each water sample. The cumulative sum of the PI and NI was computed for grading the water quality at a sampling location as m-HPI.

2.8 INFORMATION ENTROPY

In 1948, Shannon published his famous paper *A Mathematical Theory of Communication*, in which he devoted a section to what he calls Choice, Uncertainty, and Entropy. In this section, Shannon introduced an H function of the following form as shown in Eq. (2.5):

$$H(X) = -K\sum_{j=1}^{N} p(x_j)\, log[p(x_j)] \tag{2.5}$$

Where K is a positive constant. Shannon then states that "any quantity of this form, where K merely amounts to a choice of a unit of measurement, plays a central role in information theory as measures of information, choice, and uncertainty". j denotes a discrete data interval, x_j denotes the outcome corresponding to the interval j, and $p(x_j)$ denotes the probability of occurrence of the outcome x_j. The variable X can only have N possible outcomes. H assumes a maximum value if all the outcomes of the random process are equiprobable. Consequently, $H = 0$, if the probability of a certain state is 1 which indicates maximum certainty or information about the system state, and $H = \log_2 N$ when all the outcomes of the random process are equiprobable. Thus, in carrying out an experiment of N possible outcomes with probabilities $p_1, p_2, p_3, \ldots, p_N$ the above equation measures the uncertainty associated with the results of the experiment or in other terms the average amount of information obtained from the experiment. It can be intuitively concluded that the larger amount of information gained upon the knowledge of the outcome means a greater uncertainty associated prior to its occurrence. The entropy of any process/variable always assumes a positive value within limits defined as:

$$0 \leq H(X) \leq \log N \tag{2.6}$$

When the entropy is minimum, $H_{min} = 0$, the system is completely ordered and there is no uncertainty in its structure. Similarly, the maximum entropy H_{max} can be considered as a measure of maximum uncertainty or disorder.

Information entropy is a measure of the unpredictability of a random event, or equivalently the average information derived from its occurrence. It may also be defined as a measure of the uncertainty, disorder and diversity of a particular attribute. It is measured in bits or nats as per the base of the logarithm used to define the entropy. Entropy theory is multifaceted theory and has been applied to a wide spectrum of areas (Singh, 2014). A large number of random hydrological and meteorological processes existing in the environment bring into limelight the effectiveness of information theory in their applications (Cheng et al., 2004; Fleming, 2007; Mishra et al., 2009; Singh, 2013; Singh, 2014).

2.9 APPLICATION OF INFORMATION ENTROPY IN WATER QUALITY AND HYDROLOGICAL STUDIES

2.9.1 Entropy Weighted Water Quality Index (EWQI)

Entropy based weights have emerged as a useful technique which employ information entropy to assign weights to water quality parameters (Pei-Yue et al., 2010; Amiri et al., 2014). The assignment of weights to a particular parameter at a particular location is dependent on the uncertainty of its occurrence at that location. Higher uncertainty of occurrence at a location imply lower weights of parameters at that location (Amiri et al., 2014). The entropy weights consequently reduce the effect of any error in judgement arising from subjective opinion of water quality experts. Such judgements and opinions often leads to loss of valuable information about the water quality of a location.

Aggregation of weights and quality rating scale of all the parameters into a cumulatively derived numerical score is termed as the entropy weighted water quality index (EWQI).

In the development of EWQI, an initial matrix of the water quality parameters is normalized so as to eliminate any error caused by different dimensions and the ratio of the index value of a parameter in a water sample is calculated. The information entropy and the entropy weight of the parameter in that sample reveals that parameters with low entropy value corresponding to higher entropy weights have the maximum influence on the overall water quality (Gorgij et al., 2017; Islam et al., 2017). A quality rating scale of each parameter is calculated as per actual concentration divided by permissible concentration. EWQI is calculated by multiplying the quality rating scale with the entropy weight of the parameter and is classified according to the quality gradation scale.

2.9.2 Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS)

In recent years, MCDMs have shown efficient performance in the fields such as agricultural water quality assessment, water quality monitoring, reservoir operation, and biodiversity rehabilitation. TOPSIS was originally developed by Hwang and Yoon in 1981 with further developments by Yoon in 1987, and Hwang et al. in 1993. A combined application of TOPSIS with entropy weights utilizing rough set theory illustrated its flexibility in monitoring programs. There are two types of criteria in this method. Positive criteria are those that should be increased and negative ones are those which need to be decreased in order to mitigate risk. TOPSIS method based on information entropy aims in arriving at an alternative or scenario which is nearest to the positive ideal solution (PIS) and farthest from the negative ideal solution (NIS) by employing closeness coefficients (CC). Information entropy, again serves as a powerful tool in the measurement of uncertainty in a system such that when the information is specified, the uncertainty gets reduced or removed. TOPSIS performed properly in sensitivity analysis of different physico-chemical parameter weights. TOPSIS may also be utilized for validation among indices by coefficient of determination of their closeness coefficients.

2.9.3 Entropy based Diversity and Disorder Indices

The measurement of species diversity serves a fundamental role in biodiversity conservation (McIntosh, 1967; Lande, 1996). The concept of diversity index post-monsoons from the desire for both theoretical and practical reasons to gain information on the number of species found in a given region of space-time as well as on its abundancy relative to others in the same community (Maurer and McGill, 2011). A diversity index equal to 1 signifies complete evenness and a value of 0 signifies complete unevenness

(Mishra et al., 2009). Diversity indices may be variance based (Simpson's diversity index) and entropy based (Shannon's diversity index). Variance based diversity indices measure the dispersion of a data set from its mean concentration metrics. For example, if the WQI increases in one month and remains constant in all other month, variance based diversity indices will decrease faster than entropy based diversity indices suggesting greater variability of water quality. Unlike variance, entropy based diversity indices measure the dispersion of the probability density function irrespective of mean concentration metrics and may be offered as a better measure of variability or uncertainty in a dataset (Mishra et al., 2009). The principle of maximum entropy (POME) provides consistent basis to analyze water resources and geophysical processes in general. Disorder Index (DI) is useful tools in assessing the variability of random hydrological and meteorological processes and compute the difference between maximum possible entropy of a probability distribution and actual entropy of probability distribution of a dataset. Maximum possible entropies of a dataset exist during equiprobable distributions of the dataset. Such equiprobable distributions or equally apportioned states of a system indicate its uniformity where neither of the probable states of the system is dominant in influencing the outcome of its occurrence. The DI interprets the extent of this deviation as the variability in the data set. A uniform probability distribution would signify negligible variability. A high value of DI indicates higher variability. Furthermore, several studies have utilized relative entropy concept in quantifying variability with respect to uniform distribution such that if q_x represents uniform probability distribution, the relative entropy may be calculated as $D = \sum p_x \log_2(\frac{p_x}{q_x})$. However, the expression is equivalent to DI (Silva et al., 2016). The actual entropy was computed as per the probability distribution of the data set, and for any $p_j = 0$, the corresponding entropy was estimated as $\lim_{p_j \to 0^+} p_j \log_2 p_j = 0$.

2.10 RECENT LITERATURE ON THE USE OF SHANNON ENTROPY IN WATER QUALITY AND HYDROLOGICAL STUDIES

Ozkul et al. (2000) emphasized on the requirement of potential methods to delineate the efficiency and cost-effectiveness of current monitoring networks and programs. The assessment of water quality monitoring networks with the concept of entropy quantitatively measured the information produced by a network along the Mississippi river, Louisiana. A total of 12 monitoring stations run by the Louisiana Department of Environmental Quality covering a period of 27 years between 1966 and 1992 with monthly observed values of 26 water quality variables were applied in the study. The best combination of stations in a monitoring network can be selected by a procedure that foresees the minimization of transinformation among stations in the combination. The temporal design in the study proposed the extension of the existing monthly sampling intervals to bimonthly sampling frequencies at almost all of the majority sampling sites.

Mogheir et al. (2004) applied information entropy in characterization of spatial variability of synthetic ground water quality data. Discrete and analytical approaches were carried to calculate transinformation (T model), information transfer index and correlation coefficients (C model). The results of the synthetic data analyses demonstrated that the class interval and size of the data affect the T model results. Both C and T models can be employed in characterizing spatial variability. Both the analytical models fit the discrete data models well with high values of coefficient of determination.

Mishra et al. (2009) employed the entropy concept for the investigation of spatial and temporal variability of precipitation time series in the State of Texas, USA. Marginal entropy served as an important tool for investigating the variability associated with monthly, seasonal and annual time series. Apportionment and intensity entropies were also employed for investigating the intra-annual and decadal distributions of monthly and annual precipitation amounts and numbers of rainy days within a year and decade respectively. Hurst exponent and Mann Kendall test were performed to evaluate long-term persistence and trend in the variability of precipitation. A total of 43 stations having a common data period from 1925 to 2005 were chosen for investigating the precipitation variability The results indicated distinct spatial patterns in annual series and different seasons. The variability of number of rainy days as well as the precipitation amount increased from east to west of Texas. Annual series variability seemed to be having lesser disorderliness in comparison to the constituent seasonal time series. Spatial variability in monthly precipitation amount and number of rainy days indicated inconsistency in the intensity of precipitation pattern.

Karamouz et al. (2010) presented an entropy-based approach for the design of an on-line water quality monitoring network for the Karoon river, Iran. The number and locations of the sampling sites and frequencies were determined by minimizing the redundant information, which was quantified by the information theory. A pair-wise comparison of the several water quality variables used in the design of monitoring networks was used to combine the spatial and temporal frequencies calculated for each water quality variable. Data acquisition, transmission and processing were designed for the study area after selection of sampling frequencies. The results showed that the proposed approach were capable of being used in the design on on-line water quality monitoring networks.

Liu et al. (2010) combined the fuzzy mathematics method and the information entropy method to establish an improved fuzzy comprehensive evaluation method for water quality assessment. An exponential membership function was adopted to solve zero-weight problem, and the information entropy has been used to modify the coefficients of weight in order to exploit useful information data to maximum content. The results of the analysis showed that the fuzzy comprehensive evaluation method was superior than the traditional method.

Liu et al. (2012) applied information theory for data analysis and extraction of useful information for coastal water quality management from a semi-continuous water quality monitoring system installed in Yunlin Offshore Industrial Park (YOIP). The system provided real-time water quality information such as pH, water depth, dissolved oxygen, turbidity, conductivity and chlorophyll. Shannon entropy was applied so as to measure the inherent uncertainty of the water quality information. The transformation of high frequency measurements to entropy based readily available data greatly revealed the underlying patterns in the original data yielding a new insight and valuable perspectives to the original data. A spike in entropy, meant a high fluctuation in water quality indicating possible pollution by a pollution source. The Shannon entropy for turbidity reached a value of 0.8 between June 12 and June 24 and for chlorophyll it reached a value of 2 in and around June 30. Thus, Shannon entropy can be used as a supplementary indicator along with original water quality data to signify episodes of water quality degradation.

Fagbote et al. (2013) determined the WQI of the ground water of bitumen deposit impacted farm settlements using the entropy weighted method in a bitumen deposit area of Western Nigeria for the dry and rainy reasons of 2008 and 2009. A total of 15 parameters were taken into consideration for evaluating the EWQI. Conductivity, pH, turbidity, phosphate and total coliform values obtained from the analysis of 6 bore well samples were found to be well above their permissible limits. The EWQI ranked the bore wells as ‘excellent’ in the dry season of the year 2009 and just ‘good’ in the rainy season of the same year. However, the sampling stations were ranked extremely poor at least once out of the four seasons included in the study. Correlation coefficient matrices of the EWQI and the parameters showed significant relationships and high concentration of total coliform in most of the shallow wells in the environment due to bitumen deposit rendered them unfit for human consumption.

Sianaki and Masoum, (2013) demonstrated fuzzy TOPSIS decision-making approach to quantify and evaluate consumers' preferences at micro-level when using according to real-time price scheme of demand response in order to manage efficiently the use of appliances. Householder participation was maximized in such demand response programs.

Amiri et al. (2014) studied the ground water quality and its suitability for drinking in the Lenjanat Plain aquifer, Iran. Fifty-nine ground water samples from the study area were evaluated based on WHO and Iranian drinking water standards. Seasonal sampling from the selected monitoring sites was done during the period of 2009-2010. The parameters taken into consideration for the evaluation of the water quality were electrical conductivity, pH, total dissolved solids, Ca^{2+}, Na^{+}, K^{+}, Mg^{2+}, HCO_3^-, NO_3^-, SO_4^{2-}, Cl^- and F^-. An information entropy method was used to assign weights to each parameters because personal judgements by experts lead to loss of useful information about parameter weights. The results showed that over 57 and 74% of the ground water samples were in the range of "excellent" to "medium quality" based on WHO and Iranian standards respectively in the wet season and around 42 and 62% of the samples were in the range of "excellent" to "medium quality" based on WHO and Iranian standards respectively in the dry season.

Djebou et al. (2014) analyzed the effects of watershed topography on variability of summer precipitation (June-July August) in southwestern United States. The study utilized historical data over five decades from 370 meteorological stations and entropy theory to assess space-time variability. The JJA total precipitation and number of precipitation events were considered in the analysis. The influence of watershed topography on precipitation variability was investigated using spatial regionalization combining multivariate statistical techniques. The principle of maximum entropy was utilized to compute the disorder index signifying the deviation or variability of the precipitation distribution from its uniform distribution (maximum entropy). Seasonal precipitation characteristics were found to increase from east to west in the south-west US. In the east and humid parts, variability in summer precipitation characteristics were lower while the variability increased in arid regions. The study also concluded that 19-27% of the precipitation was related to terrain slope, clustered precipitation variability corresponded to terrain variation, and there is a stabilizing effect of hilly relief features on the JJA precipitation charactersitics.

Jianqin et al. (2014) evaluated a model based on the combination of principal component analysis (PCA) and Information Entropy (IE). The model combined these two methods to obtain weights of indicators and the proposed model was applied to assess the reused water quality of Jinshui river in Zhengzhou city in 2009. The results of the analysis proved the method to be feasible and practical, and it can provide a theoretical basis and decision reference for the utility of unconventional water.

Wu et al. (2015) evaluated the sensitivity of entropy weight to sample statistics in assessing water quality. The study was based on 1002 stochastic samples and their relationships with mean, standard deviation, coefficient of variation was examined. The study adopted two modes of samples to quantify the sensitivity namely the varied number of samples mode (mode 1) and the fixed number of samples mode (mode 2). The results showed that in mode 1, no significant correlations were observed between entropy weight and sample statistics while in mode 2, a significant quadratic relationship between entropy weight and mean, and a linear relationship between entropy weight and standard deviation were observed. Also, sensitivities of entropy weights to the variation of maximum and minimum of data could be expressed by quadratic equations.

Zhao et al. (2015) examined precipitation variability based on information entropy over a data set of 50 years in Xinjiang, northwestern China. On the basis of precipitation data of 53 meteorological stations from 1960 to 2008, entropy method analyzed the spatial variability over monthly, seasonal and annual timescales. The spatial distribution was affected by landscape topography in all the time scales. The non-parametric Mann Kendall test was used to analyze the changes in precipitation distributions. A precipitation concentration index was developed and variability in summer season contributed less to annual variability than of other seasons. Overall, the precipitation variability was found to increase in northern region of Xinjiang, especially mountainous regions. The precipitation concentration index had strong correlation with disorder indices, and could also be used to quantify the variability. Significant trend changes were observed in early 1980s in mountainous regions, and in the basin regions in the late 1990s.

da Silva et al. (2016) employed Shannon entropy in assessing space-time variability of rainfall and streamflow in semi-arid northeastern region of Brazil. The study utilized the principle of maximum entropy to analyze water resources along with geomorphic and land use characteristics. Mean values of marginal and relative entropies were evaluated for a ten-year period from 189 stations in the study area. Marginal entropy was utilized to evaluate randomness in the time series and relative entropy approach was utilized to analyze the similarity or dissimilarity among different sources. Results showed that uncertainty in streamflow data was higher than in rainfall data due to disorder and randomness of streamflow records. An analysis of the relative entropy depicted that precipitation and streamflow carried the same information over the year and rainy season. This could be attributed to high flow originating from high rainfall in the rainy season and low flow originating from low rainfall in dry season.

Zhang et al. (2016) employed information entropy in assessing the spatiotemporal patterns of precipitation in the Huai river basin, China. Trends of precipitation variability were quantified using modified Mann Kendall test. The study used daily precipitation data from 36 meteorological stations covering the period of 1961-2005. Increasing non-uniformity of annual precipitation and number of precipitation days was observed from south to north. Disorder indices employed exhibited variability over different temporal scales. Precipitation variability was larger at shorter time scales and smaller at shorter time scales. Significant relations were identified between disorder indices and extreme precipitation events. The study concluded that higher precipitation variability could be credited to higher frequency of extreme precipitation regimes.

Faiz et al. (2017) examined the precipitation variability and uncertainty of river flows based on diversity indices and Shannon's information entropy theory in Hindu Kush Himalayan and Karakoram river basins of Pakistan. The data sets ranged from the period of 1960-2012. The results indicated that the Shannon diversity index depicted precipitation variability on all meteorological stations as compared to Shannon's diversity index. The maximum precipitation variability was observed at Chilas station. The uncertainty analysis of stream flow depicted higher entropy values for Indus river gauged at Mangla station. Higher entropy meant greater chaos of variables and lower certainty. Rivers with high stream flow variability depicted low entropy of its distribution and therefore higher stream flow concentration in the annual cycle.

Gorgij et al. (2017) evaluated the groundwater quality for 21 groundwater samples from the Azarshahr plain in Iran using entropy theory and its results were compared with the spatial autocorrelation of effective parameters of water quality. Entropy weights were used to prevent errors arising from expert judgements and subjective opinions. The spatial autocorrelation assessment confirmed the entropy theory results. Using the entropy weighted WQI groundwater quality was classified into categories such as excellent, good, moderate, poor and extremely poor. Bicarbonate had the greatest spatial autocorrelation and the lowest entropy value while manganese had the lowest spatial autocorrelation and highest entropy value. In the study, 54% of the groundwater samples had water quality "moderately" to "extremely poor".

Islam et al. (2017) characterized the groundwater quality ranks for drinking purposes using entropy method, spatial autocorrelation and geostatistics in Sylhet district, Bangladesh. A total of 91 samples were collected from shallow, intermediate and deep tube wells. The results showed nitrate was the most influencing parameter affecting water quality followed by sulfate and arsenic. Based on EWQI classifications, 60.45% and 53.86% of water samples were having excellent to good qualities while the remaining varied from medium to extremely poor. A Gaussian semivariogram model was chosen to the best fit model, and groundwater quality indices having a weak spatial dependence suggested that both anthropogenic and geogenic played a pivotal role in spatial heterogeneity of groundwater quality.

Keum et al. (2017) reviewed the applications of information entropy in water quality monitoring design. The review summarized common entropy terms used in monitoring network designs and dealt with its applications in water monitoring design for four broad categories namely precipitation, streamflow and water level, water quality, and soil moisture and groundwater networks.

Sahoo et al. (2017) evaluated water quality with application of Bayes' rule and entropy weighted method. Six monitored parameters from gauging stations of Brahmani river were used in assessing the trends in water quality. Bayes' rule was applied for comprehensive assessment. The likelihood estimates were obtained from the normal distribution and was pre-owned for posterior probability calculation. Weights of parameters were determined using information entropy. Aggregative index evaluation method was applied. The overall water quality was better in dry seasons than in wet seasons due to less effects of leaching. The evaluation of water quality suggested that the water was acceptable for second grade surface source protection zones for centralized drinking water.

Wu et al. (2017) assessed the water quality of Shahu Lake in the semi-arid loess area of northwestern China to provide valuable information on lake water quality for decision making. The monitored parameters were compared with water quality standards of China and overall water quality assessment was done using entropy weighted water quality index based on 20 selected parameters. The lake water quality was assessed for irrigation purpose also. The study indicated that the inorganic contamination in the lake was more significant than the organic contamination. The overall water quality of the lake assessed by EWQI graded the lake "poor" and "very poor" with sulfate, total dissolved solids, total hardness and chloride being the major contributing factors. The lake water was suitable for irrigation from alkalinity point of view but unsuitable from salinity point of view. The water quality of lake also had high permeability index (class 1).

Zahedi (2017) modified conflicts between Drinking Water Quality Index (DWQI) and Irrigation Water Quality Index (IWQI) in water quality ranking of shared extraction wells using MCDMs. The study based in Karaj plain, Tehran employed three MCDM techniques such as Ordered Weighted Averaging (OWA), Compromise Programming (CP) and TOPSIS to alleviate contradiction in groundwater wells' ranks. The study reported that the application of TOPSIS and entropy weight along with the utilization of rough set theory resulted in greater reliability analysis from the sensitivity of different physicochemical parameter weights. The study also reported that the introduction of MCDMs in conflict removal was observed to a great extent in the case study and could possible reanalyze the groundwater quality classes and prioritizing extraction from groundwater resources.

Zahedi et al. (2017) introduced a novel framework of evaluating and validating groundwater quality indices using MCDMs namely TOPSIS and CP. The study based in Varamin plain, Iran used MCDMs to address shortcomings of eclipsing and exaggeration of groundwater quality indices in 92 groundwater extraction wells. The results demonstrated that TOPSIS could present a more precise analysis for classes with low number of wells.

2.11 RESEARCH GAP AND HYPOTHESIS

This chapter presented an extensive review on the environmental issues related to water quality monitoring and the difficulties faced by researchers and policy makers in extracting useful information from the observed datasets. In the recent years, there has been keen interests in identifying the pollution sources from complex datasets and incorporating them in developing several water quality models. Various MSTs have been employed for identifying probable sources for water quality variability and identification of latent pollution sources. Attempts have been made to express the water quality in terms of WQIs. Most of the conventional water quality indices are crisp and deterministic in the sense that they rely on the accurate determination of water quality characteristics and on the crisp manners of aggregation using deterministic tools that can resolve sets of characteristics into overall water quality. But all experimental methods to assess the water quality are also strongly reductionist in the sense that they try to assess the nature of the whole on the basis of few parts of the whole. This introduces elements of subjectivity and uncertainty in the endeavor at the outset. No literature suggests any methodology for ranking of various resources. Various researchers claim monitoring datasets to be “data-rich but information-poor”, however fail to provide any significant solution. Very limited literatures are available on monitoring programs aided with MSTs.

Application of information entropy offers better understanding of random processes occurring in nature. Most of the studies employing information entropy in the assessment of space-time variability has been limited to hydro-meteorological variables. New areas finding the application of entropy have since continued to unfold. The entropy theory is indeed versatile and its application is widespread.

Various hypotheses pertaining to the water quality monitoring programs can be proposed:

- Based on the various methodologies developed so far, can we define various latent pollution sources for the water bodies in the north-east India as these areas have been untouched so far?
- Can the information entropy be applied on random water quality datasets for expressing the water quality of a particular water body as successfully as it has been used for understanding the random hydro-meteorological processes occurring in nature?
- Can a general method be developed for ranking of water bodies for better allocation of resources?
- Will it be possible to develop a suitable methodology for identification of ideal sampling sites?

The primary objectives of the present study will be aimed at answering these queries. This study will be a mixture of mathematical and experimental evidences. The results of the experimental investigation coupled with that of mathematical models will contribute to the refinement of the research questions, methodologies and theory.

3

Chapter

Materials & Methods

This chapter covers

- Design of research work
- Study area
- Sampling strategies and analytical procedures
- Methodologies for PCA, CA, EWQI, ADI, Hazard index, TOPSIS, Disorder index, Diversity index

---

3.1 DESIGN OF RESEARCH WORK

In order to accomplish the objectives of the present study, research work was carried out in four phases. Phase 1 was associated with survey of study area, identification of sampling locations, collection of water samples and analysis of collected samples in laboratory. In Phase 2, multivariate statistical techniques (MSTs) were applied on the observed dataset to discriminate the sources of variation of water quality. Dataset was treated with cluster analysis (CA), discriminant analysis (DA) and principal component analysis (PCA) to identify the latent pollution sources. Phase 3 was associated with application of information entropy for evaluation of water quality and overall ranking of sampling sites. Entropy weighted irrigation water quality index (EIWQI) and entropy weighted heavy metal contamination index (EHCI) have been proposed. Risk assessment due to heavy metals ingestion was also included in this phase. In Phase 4, entropy based disorder index and diversity index were used to identify the ideal sampling location for monitoring of water bodies. Flow chart of research work is shown in Fig. 3.1.

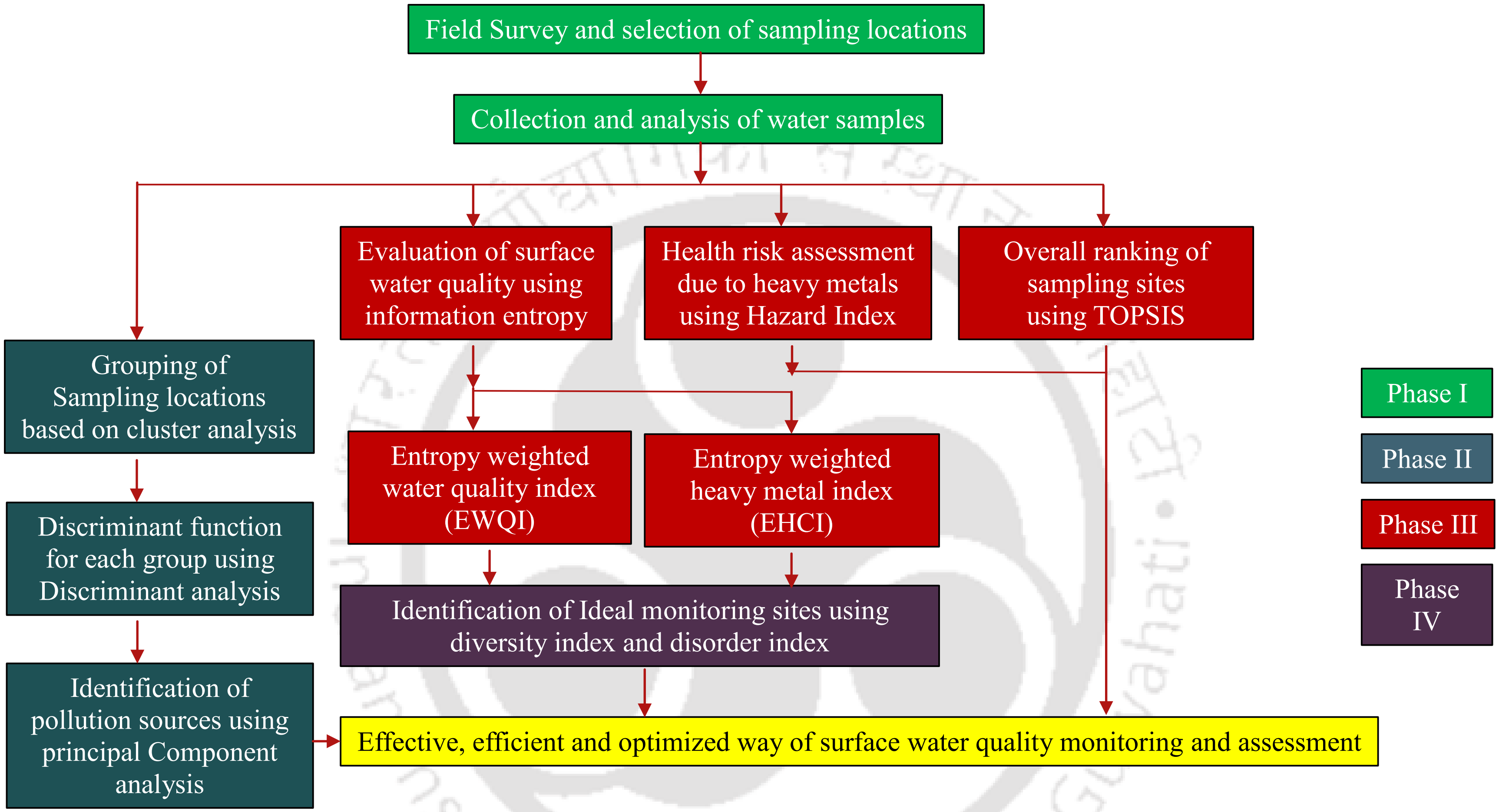


Fig. 3.1. Design of research work

PHASE 1: RECONNAISANCE SURVEY AND DATA ACQUISITION

3.2 STUDY AREA (BRAHMAPUTRA RIVER AND TRIBUTARIES)

The study areas comprise of Brahmaputra River and its tributaries (Baralia, Puthimari, Pagladia, Beki, Manas, Kolong and Kameng Rivers). Brahmaputra River originates from Himalayan Lake Manasarover in Tibet and flows through China, India and Bangladesh. It enters India in Arunachal Pradesh and then flows down to the plains of Assam. Several tributaries join Brahmaputra on northern as well as southern banks. All the tributaries considered in the present study lie to the northern bank except Kolong. Puthimari River originates from the foothills of the Himalayan Ranges in Bhutan. After crossing the Indo-Bhutan border, it bifurcates into two separate rivers (Baralia and Puthimari) near Bornodi Widlife sanctuary Arangajuli, Assam and poses all the characteristics of a flashy river. It meanders freely and has many loops, the slope being somewhat flatter in lower reaches. Baralia River flows through the heart of Rangia, a town in Kamrup rural district of Assam, whereas Puthimari River flows through the periphery of the city. Both the rivers meet near Hazo town before joining Brahmaputra. The approximate lengths of Baralia and Puthimari is 39.1 and 139 km respectively. The region is characterized by heavy floods due to intense rainfall during monsoon. Pagladia originates from southern slopes of Bhutan Hills in the form of two streams Pagla and Dia that meet near Chowki. It flows through Bagsa and Nalbari before joining Brahmaputra near Sotemari. The river flows for a length of 19 km in the hilly track of Bhutan and the rest 177.80 km flows through Assam. It is responsible for frequent floods in the north bank of the Brahmaputra in Assam. Beki River flows down from the Bhutan region but a large portion flows in Assam. In Bhutan, the river is known as Kurissu River. In Assam, the river flows through Barpeta district (Kalita et al., 2015). The river and its drainage channels flow through the plains of Assam for about 85 km. Manas river has its origin in the Himlayan foothills between southern Bhutan and India. It is regarded as the largest river system of Bhutan and debouches into India through western Assam, covering a distance of 104 km before confluence into the Brahmaputra river at Jogihopa. The major tributary of Manas River in Assam is Aie River which joins it in Bangpari. The Manas River flows through the outskirts of Bongaigaon, which is a major city in the state of Assam. The climate is humid and sub-tropical in the region and it experiences the highest rainfall during the month of June and July (more than 350mm). It receives less rainfall during winter (November to February). The average temperature of the city varies from 10 to 35 ˚C. Kolong River is one of the most polluted rivers of Assam. The tributary is about 250 km long and spreads in Nagaon, Morigaon and Kamrup districts. It diverts out from the Brahamputra near Jakhalabandha and meets it again at near Guwahati. On its way, various minor streams (Diju, Misa, Haria and Digaru) join, thereby making the stormwater runoff, base flow from ground water seepage and waste water discharges from various towns contribute to its flow. Rainfall distribution in the Kolong River basin follows a typical monsoon pattern with peak rainfall during monsoon and scanty precipitation during winter. Kameng River (Jia Bhoreli in Assam) originates from the glacial lake (Tawang, Arunachal Pradesh) and joins Brahmaputra at Tezpur, the centre of administration for the Sonitpur district (Assam). The total length of the river is approximately 264 km and its drainage basin is about 11,843 km^2. Dirang Chhu, Bichom, Tenga river and Tippi are its few major tributaries (Khound and Bhattacharyya, 2017). Industrial activities are very less along the stretch but river receives untreated or partially treated sewage and surface runoff. Major agricultural activities in the catchment area are tea plantation and rice cultivation (Khound and Bhattacharyya, 2017).

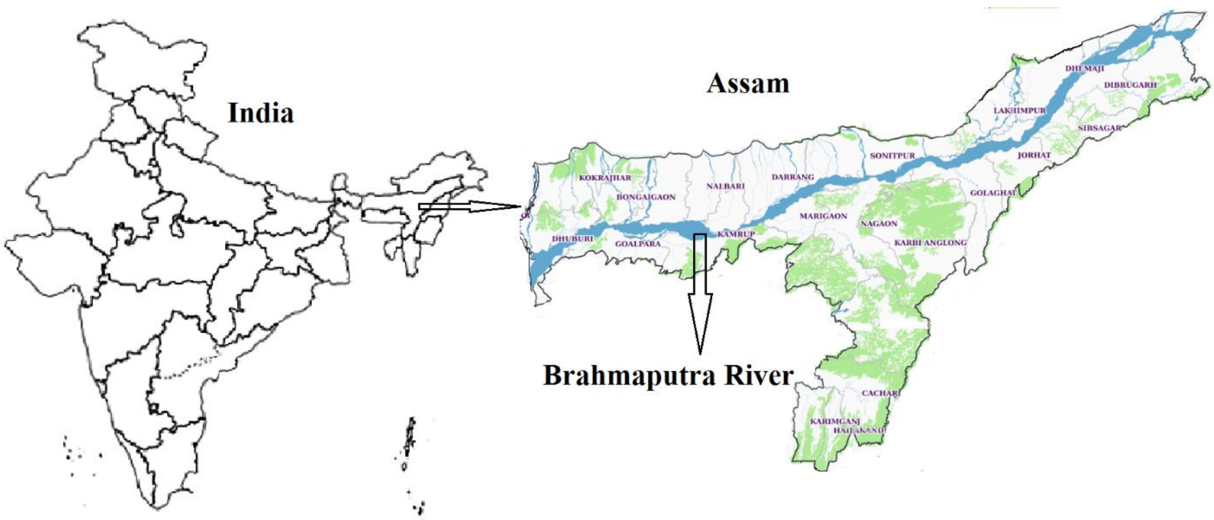
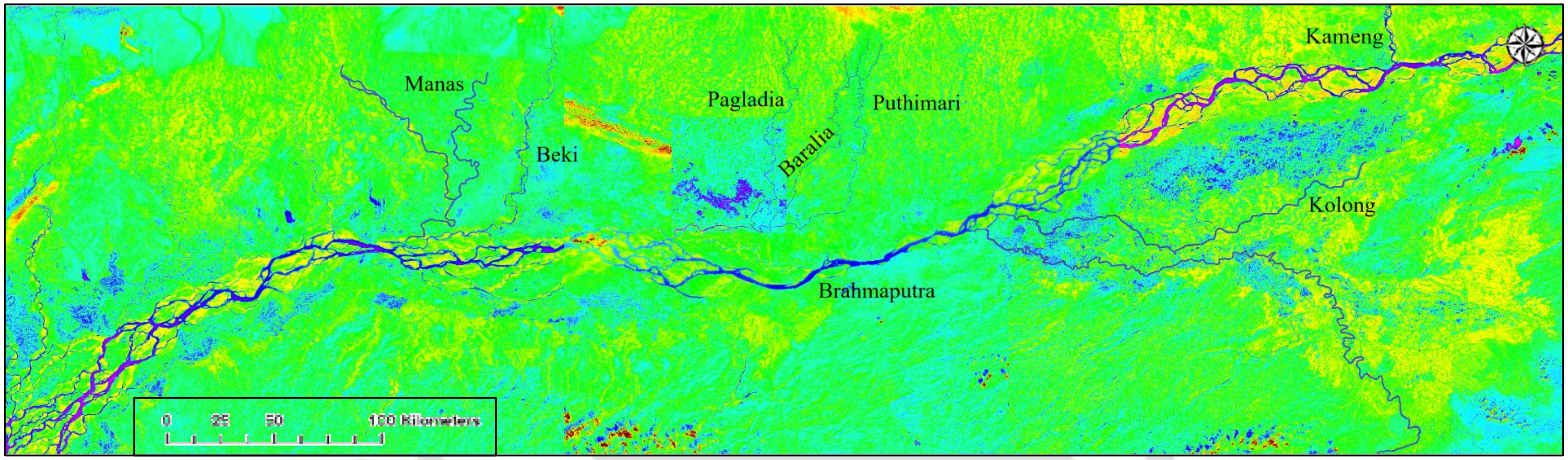


Fig. 3.2 Map of Assam



Baralia River

LOCATION	LATITUDE	LONGITUDE
SPBR1	26°37'47.46"N	91°38'10.98"E
SPBR2	26°35'23.46"N	91°38'38.52"E
SPBR3	26°32'52.20"N	91°37'7.50"E
SPBR4	26°26'15.27"N	91°37'28.18"E
SPBR5	26°26'11.16"N	91°37'18.60"E
SPBR6	26°26'2.49"N	91°37'12.58"E
SPBR7	26°25'30.83"N	91°35'18.29"E
SPBR8	26°21'48.91"N	91°31'27.83"E
SPBR9	26°17'51.23"N	91°29'43.91"E

Puthimari River

LOCATION	LATITUDE	LONGITUDE
SPPR1	26°37'29.64"N	91°39'50.22"E
SPPR2	26°31'24.67"N	91°40'20.33"E
SPPR3	26°22'2.52"N	91°39'13.86"E
SPPR4	26°19'38.23"N	91°38'42.54"E
SPPR5	26°17'57.48"N	91°38'15.24"E
SPPR6	26°15'56.95"N	91°30'18.57"E

Kolong River

LOCATION	LATITUDE	LONGITUDE
SPKR1	26°21'7.27"N	92°42'52.30"E
SPKR2	26°21'38.71"N	92°42'28.70"E
SPKR3	26°21'29.02"N	92°41'16.61"E
SPKR4	26°20'58.81"N	92°40'47.36"E
SPKR5	26°20'29.73"N	92°40'41.65"E
SPKR6	26°19'38.79"N	92°40'25.94"E
SPKR7	26°18'27.24"N	92°39'47.66"E
SPKR8	26°17'27.10"N	92°38'49.18"E
SPKR9	26°15'44.21"N	92°34'51.07"E
SPKR10	26°13'59.13"N	92°32'10.54"E

Pagladia River

LOCATION	LATITUDE	LONGITUDE
SPPaR1	26°31'16.47"N	91°30'29.73"E
SPPaR2	26°29'8.04"N	91°28'10.10"E
SPPaR3	26°27'58.34"N	91°27'40.00"E
SPPaR4	26°27'34.46"N	91°27'16.37"E
SPPaR5	26°27'2.09"N	91°27'36.66"E
SPPaR6	26°25'48.40"N	91°28'31.58"E
SPPaR7	26°24'24.73"N	91°28'38.39"E
SPPaR8	26°22'54.15"N	91°27'49.05"E
SPPaR9	26°20'9.03"N	91°28'38.07"E

Beki River

LOCATION	LATITUDE	LONGITUDE
SSBeK1	26°35'51.13"N	90°57'59.25"E
SSBeK2	26°33'6.97"N	90°58'18.60"E
SSBeK3	26°32'31.14"N	90°56'13.42"E
SSBeK4	26°31'11.42"N	90°55'26.69"E
SSBeK5	26°29'39.65"N	90°55'4.13"E
SSBeK6	26°26'25.24"N	90°55'1.90"E
SSBeK7	26°23'17.19"N	90°54'51.63"E
SSBeK8	26°19'14.24"N	90°52'33.84"E

Manas River

LOCATION	LATITUDE	LONGITUDE
SSMR1	26°38'3.20"N	90°29'53.68"E
SSMR2	26°35'31.44"N	90°30'52.83"E
SSMR3	26°34'9.88"N	90°33'6.68"E
SSMR4	26°32'16.44"N	90°35'43.88"E
SSMR5	26°30'37.89"N	90°37'21.70"E
SSMR6	26°30'0.08"N	90°39'11.53"E
SSMR7	26°27'53.07"N	90°40'14.26"E
SSMR8	26°25'40.48"N	90°41'47.51"E
SSMR9	26°20'10.63"N	90°44'35.96"E

Kameng River

LOCATION	LATITUDE	LONGITUDE
SSKaR1	26°58'20.52"N	92°48'3.28"E
SSKaR2	26°56'48.18"N	92°49'30.27"E
SSKaR3	26°55'2.77"N	92°50'57.59"E
SSKaR4	26°51'59.42"N	92°51'48.12"E
SSKaR5	26°49'26.15"N	92°51'45.79"E
SSKaR6	26°47'58.36"N	92°52'57.11"E
SSKaR7	26°46'2.80"N	92°52'7.57"E
SSKaR8	26°43'6.03"N	92°51'14.18"E
SSKaR9	26°40'16.75"N	92°51'53.90"E

Brahmaputra River

LOCATION	LATITUDE	LONGITUDE
SS1	26˚11'54.7" N	91˚45'25.4" E
SS2	26˚11'21.86" N	91˚44'26.6" E
SS3	26˚10'32.8" N	91˚43'49" E
SS4	26˚10'25.5" N	91˚43'41" E
SS5	26˚10'20" N	91˚43'31.6" E
SS6	26˚10'15.5" N	91˚42'59.4" E
SS7	26˚10'20.2" N	91˚41'3.7" E
SS8	26˚10'18.2" N	91˚40'39.5" E
SS9	26˚10'17" N	91˚40'24.19" E

Fig. 3.3 Study area with GPS coordinates of sampling locations

3.3 SAMPLING STRATEGIES AND ANALYSIS

3.3.1 Sampling strategies

A preliminary survey work for this study was done before sample collection. It was essential to make a recce survey of the study area during the planning stage, noting all sources of wastes, all entering tributaries that might contribute a potential pollutant threat, and all uses and abstractions of the water. Survey also encompassed the collection of background information such as geography, topography, climate and weather, hydrology, hydrogeology, land use, urbanization, industrialization and agriculture etc. in catchment area. Based on information, suitable sampling sits and also appropriate number of sampling locations were decided.

3.3.2 Collection of water samples

Water samples were collected from well-mixed section of the rivers and lakes using a weighted bottle sampler. One-liter capacity, cleaned bottles, free from dust and dirt, were used for collection of water samples. Before sampling, bottles were rinsed 2-3 times with the water to be sampled. All bottles reached at the sampling site in a fully cleaned state, protected from accidental contamination. Same bottles were used only for identical selected parameters. To avoid contamination by the hands, plastic gloves were used. The sample bottles were labelled with site number and date. Collected samples were both preservative and non-preservative. In preservative samples nitric acid of 2mL/L were added to make pH less than or equal to 2 to avoid the precipitation of metals and adsorption to container wall (APHA, 2012). Preservative samples were used for analysis of heavy metals and non-preservative samples were used for other water quality parameters. After collection of samples in the field, bottles placed in an insulated cool box together with ice. After reaching in the laboratory in the samples were immediately transferred in refrigerator.

3.3.3 Analytical Procedures

Analytical procedures of Standard Methods for the Examinations of Water and Wastewaters 20th edition, published by APHA (2012) have been followed throughout the analysis. A quality control procedure was maintained throughout, including recalibration of instruments. Reagents were prepared as recommended by Standard Methods. All chemicals and reagents used in the analyses were of analytical grade unless otherwise stated. Deionized water was used for all dilutions. Standard solutions were prepared by diluting the stock solutions. Water quality parameters associated with their units, abbreviations and analytical methods used in this study are summarized in Table 3.1.

Table 3.1. Water quality parameters associated with their units, abbreviations and analytical methods used in this study

Parameters	Unit	Abbr	Analytical methods	Model
Temperature	°C	T	Thermometer	Digital Thermometer (Thomas Scientific)
pH	-	pH	pH-meter	µ pH System 361 (Systronics)
Dissolved Oxygen	mg/L	DO	DO meter	HQ30D Portable Dissolved Oxygen Meter (Hach)
Alkalinity	mg/L	TA	Titrimetric	
Hardness	mg/L	TH		
Total Solids	mg/L	TS	Gravimetric	
Total Dissolved Solids	mg/L	TDS		
Total Suspended Solids	mg/L	TSS		
Electrical Conductivity	µS/cm	EC	Electrometric	MT-112TDS) (Manti Lab Solutions)
Sodium	mg/L	Na^+	Flame photometer	µ Controller Based Flame photometer with Compressor (Type 128) (Systronics)
Potassium	Mg/L	K^+		
Calcium	mg/L	Ca^{2+}		
Fluoride	mg/L	F^-	Ion Chromatography	Metrohm 792 Basic IC
Chloride	mg/L	Cl^-		
Sulfate	mg/L	SO_4^{2-}		
Magnesium	mg/L	Mg^{2+}	Atomic absorption spectroscopy	iCE 3000 SERIES (Thermo Scientific)
Iron	mg/L	Fe		
Manganese	mg/L	Mn		
Lead	mg/L	Pb		
Copper	mg/L	Cu		
Chromium	mg/L	Cr		
Zinc	mg/L	Zn		

PHASE 2: IDENTIFICATION OF LATENT POLLUTION SOURCES

3.4 MULTIVARIATE STATISTICAL TECHNIQUES (MSTs)

Observed water quality data-sets were analysed using three multivariate statistical methods: CA, DA and PCA. CA was used to group the sampling sites into homogeneous and distinct clusters. In this study hierarchical cluster analysis (HCA) was performed for grouping of sampling sites and results were illustrated by dendrogram (tree diagram). Observed variables in this study were in different scales and units, so data-sets were treated by z-transformation. HCA was accomplished on the normalized dataset (z score) by means of the Ward's method, using squared Euclidean distances as a measure of similarity.

DA was used in this study to describe the relationships among the clusters resulted from CA. The canonical discriminant functions (CDFs) of the discriminating variables were used to discriminate among groups. DA was performed with the same raw data set comprising all water quality parameters after grouping into different clusters. The sites (clustered) were the grouping (dependent) variable, while all the measured parameters constituted the independent variables.

PCA was performed on the normalized data sets (zero mean and unit variance) separately for the all major clusters obtained from CA to compare the compositional pattern between analyzed water samples and identify the factors influencing each one. PCA extracted significant PCs and to further reduce the contribution of variables with minor significance; these PCs were subjected to varimax rotation. PCA involves following major steps (Ouyang et al., 2006):

(a) Standardization of all water quality
(b) Calculation of covariance matrix
(c) Estimation of eigenvalues and eigenvectors
(d) Development of factor loading matrix
(e) Perform varimax rotation

All statistical analysis was performed using IBM SPSS Statics 20 software. IBM SPSS is a statistical software package used for statistical analysis. Initially SPSS, stood for "Statistical Packages for the Social Sciences" but it was later changed to "Statistics Product and Service Solutions" because of its acceptance among the users of other communities. This software was developed by SPSS Inc. and acquired by IBM in 2009. The graphical user interface of IBM SPSS is written in Java.

PHASE 3: SURFACE WATER QUALITY EVALUATION EMPLOYING SHANNON ENTROPY

3.5 ENTROPY WEIGHTED WATER QUALITY INDEX (EWQI)

An entropy weighted water quality index (EWQI) is an improvement over the existing traditional WQIs which are otherwise based on assignment of weights to parameters on the basis of personal judgements and expert opinion (Amiri et al., 2014; Fagbote et al., 2014). Such judgements and opinions often leads to loss of valuable information about the water quality of a location. Steps involved in calculation of EWQI are as follows: (Li et al., (2010)

- Step 1. A matrix was developed with all 'm' water samples (m = 1, 2…m) and 'n' measured parameters (n = 1, 2…n)

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & \cdots & x_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{m1} & x_{m2} & \cdots & \cdots & x_{mn} \end{bmatrix} \tag{3.1}$$

- Step 2. Initial matrix was converted the standard grade matrix Y, to remove the error caused by different dimensions and units. After transformation, the standard grade matrix Y was obtained

$$y_{ij} = \frac{x_{mn} - (x_{mn})min}{(x_{mn})\max - (x_{mn})min} \tag{3.2}$$

where Y_{ij} is construction function of normalization for an evaluated parameter (n) in a particular water sample (m).

$$Y = \begin{bmatrix} y_{11} & y_{12} & \cdots & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & \cdots & y_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{m1} & y_{m2} & \cdots & \cdots & y_{mn} \end{bmatrix} \tag{3.3}$$

where Y = standard grade matrix

- Step 3. The information entropy was calculated by the formula:

$$E_j = (1/\ln m) \sum_{i=1}^{m} \mathrm{P_{ij}} \ln P_{\mathrm{ij}} \tag{3.4}$$

where

$$P_{ij} = \frac{y_{ij}}{\sum y_{ij}} \quad (3.5)$$

- Step 4. Entropy weight of the parameter (j) was calculated by:

$$W_n = (1 - E_j) / \sum_{j=1}^{n} (1 - E_j) \quad (3.6)$$

- Step 5. Quality rating scale for each parameter was assigned by:

$$Q_j = \left(\frac{C_j}{S_j}\right) * 100 \quad (3.7)$$

- Step 6. Calculation of WQI using the following formula:

$$WQI = \sum_{j=1}^{n} W_j\, Q_j \quad (3.8)$$

where,

Q_j = quality rating scale,

C_j = measured concentration of the parameter, and

Once the overall EWQI score is known, it can be compared against the following scale suggested by Wu et al., (2011) to determine how healthy the water is on a given period of time. The EWQI ranges, which are given below, also have been defined as:

Table 3.2. Water Quality Scale for EWQI. (Wu et al., (2011))

Rank	EWQI	Water Quality
1	< 50	Excellent
2	50-100	Good
3	100-150	Average
4	150-200	Poor
5	>200	Extremely poor

Water supplies with ratings falling in the good or excellent range would able to support a high diversity of aquatic life. In addition, the water would also be suitable for all forms of recreation, including those involving direct contact with the water.

3.6 RISK ASSESSMENT TO HUMAN HEALTH

The risk characterization and assessment of heavy metals on human health included the two major pathways- ingestion and dermal absorption. The enumeration of potential non-carcinogenic risks was done by the means of Hazard Quotient (HQ), estimated by comparing analytical values of each heavy metal with their corresponding reference dose (R_fD). Reference dose of all the observed heavy metals are given in Table 3.3. A HQ greater than 1 indicates potential concern associated with metals coming from that pathway and the summation of HQs from individual pathway result in the Hazard Index (HI) which evaluates the total potential ($HQ_{ingestion}$ + HQ_{dermal}) of non-carcinogenic risk from the water source. HQ can be calculated as:

$$Hazard\ Quotient\ (HQ) = \frac{ADD}{R_f D} \tag{3.9}$$

Average daily intake or dose (ADD) of heavy metals through any individual pathway is given by US EPA (2004) as:

$$ADD_i(in\ \mu g/kg/day) = \frac{C_0 \times I_R \times E_F \times ED}{B_w \times A_T} \tag{3.10}$$

$$ADD_d(in\ \mu g/kg/day) = \frac{C_0 \times S_A \times K_p \times ET \times E_F \times ED \times 10^{-3}}{B_w \times A_T} \tag{3.11}$$

where, ADD_i represents ADD ingestion and ADD_d represents ADD dermal. C_o is the observed concentration of heavy metal. The tabular description of the parameters has been given in Table 3.4.

Table 3.3. Reference dose for heavy metals

Metal	R_fD (μg/kg/day)
Fe	700
Mn	24
Cr	3
Pb	1.4
Cu	40
Zn	300

Table 3.4. Assumption parameters to derive average intake values

Exposure Parameters	Description	Value	References
B_W	Body weight expressed in kg	52 for average Indian man	Jain et al., (1995); Dang et al., (1996)
E_F	Exposure frequency in days/year	350	US EPA, 2004
ED	Exposure duration in years	24	US EPA, 2004
S_A	Skin surface area in cm^2	18000	US EPA, 2004
A_T	Averaging time	ED×365 for non-carcinogenic risks	US DoE, 2011
I_R	Ingestion Rate in l/day	4.05	Dang et al., (1994)
ET	Exposure time in h/day	0.6	US EPA, 2004
K_p	Dermal permeability coefficient in cm/hr	Varies for each metal	US EPA, 2004

3.7 IRRIGATION WATER QUALITY PARAMETERS

Water quality for irrigation suitability was evaluated by using permeability index (PI), Kelley's ratio (KR), magnesium adsorption ratio (MgR), sodium adsorption ratio (SAR), soluble sodium percentage (SSP), and residual sodium carbonate (RSC). Calculation of each terms are represented in equations (Eqs. 12–17):

$$PI = \frac{Na^{+} + \sqrt{HCO_3^{-}}}{Na^{+} + Ca^{2+} + Mg^{2+}} \times 100 \tag{3.12}$$

$$KR = \frac{Na^{+}}{Ca^{2+} + Mg^{2+}} \tag{3.13}$$

$$MgR = \frac{Mg^{2+}}{Ca^{2+} + Mg^{2+}} \tag{3.14}$$

$$\text{SAR} = \frac{\text{Na}^+}{\sqrt{\frac{(\text{Ca}^{2+} + \text{Mg}^{2+})}{2}}} \tag{3.15}$$

$$\text{SSP} = \frac{\text{Na}^+ + \text{K}^+}{\text{Na}^+ + \text{K}^+ + \text{Ca}^{2+} + \text{Mg}^{2+}} \times 100 \tag{3.16}$$

$$\text{RSC} = [\text{HCO}_3{}^- + \text{CO}_3{}^{2-}] - [\text{Ca}^{2+} + \text{Mg}^{2+}] \tag{3.17}$$

Concentration of all parameters used in equations (Eqs. 3.12–3.17) are meq/L.

3.8 TOPSIS

The TOPSIS model was implemented in the following manner:

Step 1. The "alternatives" (sampling locations) and the "criteria" (parameters) were specified for both the rivers to which the ranking was to be allocated according to their contamination status. Assuming there are m possible alternatives called $A = \{A_1, \ldots, A_m\}$ which are to be evaluated against "c" criteria $C = \{C_1, \ldots, C_c\}$

Step 2. The ratings to the criteria and alternatives were assigned using matrix X where x_{ij} indicated the value of alternative A_i for criterion C_j

$$X_{m\times c} = \begin{bmatrix} x_{11} & x_{12} \ldots. & x_{1c} \\ \vdots & \ldots \;\; .x_{ij} & \vdots \\ x_{m1} & \cdots & x_{mc} \end{bmatrix} \tag{3.18}$$

Step 3. The weight of the criteria was calculated on the basis of information entropy techniques as per the following equations:

$$q_{ij} = \frac{x_{ij}}{x_{1j} + \cdots + x_{mj}} \;;\;\; \forall\, j \in \{1, \ldots., c\} \tag{3.19}$$

And,

$$E_j = -\frac{1}{\ln m} \sum_{i=1}^{m} q_{ij} \ln q_{ij};\; \forall\, j \in \{1, \ldots., c\} \tag{3.20}$$

Where $0 \le E_j \le 1$ where index with higher entropy has greater variation. Therefore, weight of the criteria may be calculated as:

$$w_j = \frac{d_j}{d_1 + \cdots + d_c} \tag{3.21}$$

And, $d_j = 1 - E_j$. All the weights were aggregated to a matrix $w_{c\times c}$

Step 5. A normalized decision matrix was constructed ($N_{m\times c}$) using vector normalization method as follows:

$$r_{ij} = \frac{x_{ij}}{\sqrt{x_{ij}^2 + \cdots x_{mj}^2}} \tag{3.22}$$

Thus, $N_{m\times c} = [r_{ij}]_{m\times c}$

Step 6. A weighted normalized decision matrix was constructed (V) as follows:

$$V = N_{m\times c} \times w_{c\times c} \tag{3.23}$$

Step 7. The PIS and the NIS of the alternatives were computed as:

$$PIS = \{\max v_{ij} | v_{ij} \in V\} = (v_1^+, \dots v_c^+) \tag{3.24}$$

$$NIS = \{\min v_{ij} | v_{ij} \in V\} = (v_1^-, \dots v_c^-) \tag{3.25}$$

Step 8. The Eucledian distance of each alternative from the PIS (d_i^+) and NIS (d_i^-) were calculated as:

$$d_i^+ = \sqrt{\sum_{j=1}^{c} (v_{ij} - v_j^+)^2} \tag{3.26}$$

$$d_i^- = \sqrt{\sum_{j=1}^{c} (v_{ij} - v_j^-)^2} \tag{3.27}$$

Step 9. The closeness coefficient (CC) of each alternative was computed as:

$$CC_i^+ = \frac{d_i^-}{d_i^- + d_i^+} \tag{3.28}$$

Step 10. The alternatives were finally ranked according to their closeness coefficients.

PHASE 4: IDENTIFICATION OF IDEAL MONITORING LOCATIONS

3.9 STUDY AREA (DEEPOR BEEL)

Deepor Beel (Beel means wetland or large aquatic body in Assamese language), a freshwater lake of international importance is located at longitude 91° 35′ E to 91° 43′ E, and latitude 26° 05′ to 26° 11′ N (Choudhury and Gupta, 2017) at an elevation of 55m above mean sea level (Fig. 3.4). In November 2002, it was included in the list of Ramsar Site for undertaking preservation measures on the basis of its biological and environmental importance. Initially, the Deepor beel was linked with the Brahmaputra River. Due to various construction activities, main link has already been disrupted and it remains as a small secondary channel. The present water sources of the Deepor beel are from the catchments area and river Basishtha. The beel is a natural habitat to many varieties of birds and provide its natural resources for the livelihood of fourteen indigenous villages located in its surroundings.

The wetland has a meso-thermal climate with an average annual rainfall of 3000-4000 mm which varies within a temperature range of 10 °C – 30 °C. The wetland boasts of rich floral and faunal diversity with more than 200 species of birds out of which more than 50 species are migratory. The major inlet channels of the Beel are the Bharalu and Basistha-Bahini rivers which carry storm-water as well as sewage to the Beel. Apart from the major inlets, a large amount of surface run-off and effluents gain entry to the Beel from nearby anthropogenic influences such as Guwahati Municipal Corporation dumpsite, small industries and brick kilns located adjacent to the wetland. However, in recent years due to inflow of untreated sewage, inorganic/organic effluents and agricultural wastes the Beel has become a prey to eutrophication.

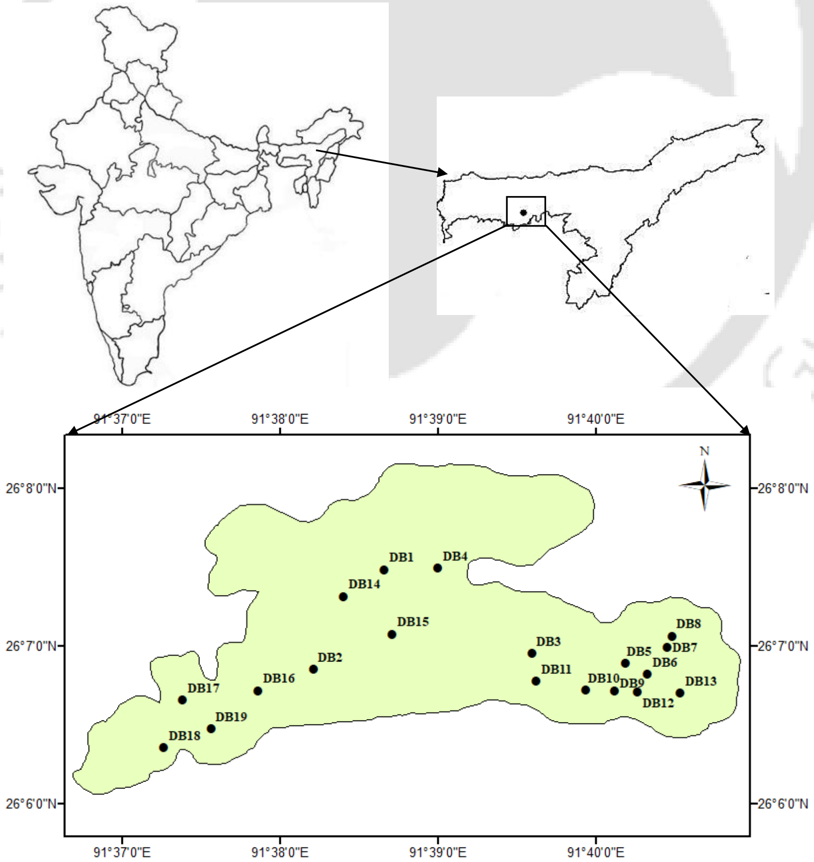


Fig. 3.4. Map of Deepor Beel

3.10 SHANNON'S DIVERSITY INDEX

The Shannon's diversity index can be computed by the aid of Shannon's entropy as an estimate of diversity in a community. The diversity index is expressed as:

$$H = -\sum p_i ln p_i \tag{3.29}$$

For a number of species N, it reaches a maximum value of ln N.

3.11 ENTROPY BASED DISORDER INDICES

Disorder index (DI) represents the deviation of the discrete marginal entropy of the data set from the maximum possible entropy of the data set and is given as:

$$\begin{aligned} &\text{Disorder index (DI)} \\ &= (\text{Maximum possible entropy under uniform probability distribution}) \\ &- (\text{Actual entropy of the data set over a time} - \text{series scale}) \end{aligned} \tag{3.30}$$

The DI interprets the extent of this deviation as the variability in the data set. A uniform probability distribution would signify negligible variability. A high value of DI indicates higher variability. The DI which is the difference between the maximum entropy value and the actual entropy value in a sampling station is computed as:

$$DI = H_{max} - H = \log_2 N + \sum_{i=1}^{n} p_i \log_2 p_i \tag{3.31}$$

where H represents the actual entropy of a data set N represents the length of the data set and p_i represent the probability of occurrence of the i^{th} event.

4

Chapter

Reconnaissance Survey and Data acquisition

This chapter covers

- Statistical summary of observed water quality data
- General characteristics of water quality parameters
- Heavy metals and associated risk assessment
- Concluding remark of Phase 1

4.1 STATISTICAL SUMMARY OF OBSERVED WATER QUALITY DATA

Water samples were collected from 68 sampling sites (9 sampling sites of Brahmaputra River and 59 sampling sites of its tributaries). A total of 22 physico-chemical parameters including heavy metals were analysed. Datasets of the analysed water quality parameters along with their statistical summary have been described in Table 4.1 (a–h) and Table 4.2 (a–h). Statistical analysis consists of maximum value (Max), minimum value (Min), average value (Mean), standard deviation (SD), skewness, kurtosis and coefficient of variance (CoV). Average value indicates the arithmetic mean of datasets. SD measures the dispersion or variation of datasets relative to its mean. A low value of SD indicates that the data points are close to the its mean. Skewness represents the asymmetry of the probability distribution of a real-valued random variable about its mean. Kurtosis is a measure of whether the data are heavy-tailed or light-tailed relative to a normal distribution. CoV represents the ratio of the SD to the mean and explains the dispersion of data points in a data series around the mean.

4.2 GENERAL CHARACTERISTICS OF PHYSICO-CHEMICAL PARAMETERS

pH is a measurement of the potential activity of H^+ and is expressed as the reciprocal of the logarithm of the H^+ activity. pH of water determines the solubility and biological availability of nutrients and heavy metals. Agricultural runoff and acidic mine drainage are the major factors that influence the pH of surface water. Low pH cause fish kill by stressing animal systems and causing physical damage. Living organisms, especially aquatic life, function best in a pH range of 6.0 to 9.0. As per BIS guidelines of drinking water (IS 10500:2012), acceptable limit of pH is 6.5-8.5. No health-based guideline value is proposed by WHO for pH. pH of all sampling sites was in the range of 7.0-8.0 except Brahmaputra and Kolong River. pH of Brahmaputra River was in the range of 7.15-8.26 and for Kolong River it varied from 6.80-7.59.

Total suspended solids (TSS) and total dissolved solids (TDS) represent non-filterable and filterable residue, respectively. TSS includes all floating materials in the water such as silt, clay, fine particles of organic and inorganic matter, plankton and algae. High flow rates, soil erosion, urban runoff, wastewater and septic system effluent, decaying plants and animals are major factors influencing the concentration of TSS in the water body. High TSS reduce the amount of light passing through the water, slow down the photosynthesis, reduce the DO level and decrease the water clarity. TDS in streams originate from natural sources, urban and agricultural run-off, sewage, and industrial wastewater. There is no health hazard associated with the ingestion of TDS in drinking water. However certain components of TDS such as Cl^-, SO_4^{2-}, Ca^{2+}, Mg^{2+}, CO_3^{2+} may affect corrosion or encrustation in water-distribution systems. As per BIS guidelines of drinking water, acceptable limit of TDS is 500 mg/L. No health-based guideline value is proposed by WHO for TDS. Average concentration of TSS and TDS at all sampling sites were below 250 mg/L except Kameng River. In Kameng River TDS was in the range of 70-605 mg/L.

Alkalinity is a measure of the capacity of water to neutralize acids. Alkaline compounds in the water such as HCO_3^-, CO_3^{2-}, and OH^- remove H^+ ions and reduce the acidity of the water. It acts as a buffer and helps to keep the water's pH stabilized. Major natural sources of alkalinity include rocks which contain CO_3^{2-}, HCO_3^-, and OH^- compounds. Some other anions such as BO_3^{-3}, SiO_4^{4-} , and PO^{3-} also impart alkalinity in water bodies. Certain plant activities and discharge of industrial wastewater are the other factors which influence the alkalinity. Alkalinity is important in determining a stream's ability to neutralize acidic pollution from various natural and anthropogenic sources. It's one of the best measures of the sensitivity of the stream to acid inputs. It is also important for fish and aquatic life because it protects or buffers against rapid pH changes. As per BIS guidelines of drinking water, acceptable limit of total alkalinity (TA) is 200 mg/L as $CaCO_3$. TA of all sampling sites was well below the acceptable limit.

Total hardness (TH) is defined as the concentrations of multivalent cations (mostly Ca^{2+} and Mg^{2+}) present in water. The major natural sources of TH in water are dissolved metallic ions from sedimentary rocks, seepage, and runoff from soils. Hard water has no harmful effect on human health but causes serious problems in industrial equipment/components like cooling towers, boilers and other equipment that contains or processes water. High TH can result in abnormal cloudiness and the formation of scale. In domestic use, hard water requiring considerably more soap to produce a lather. As per BIS guidelines of drinking water, acceptable limit of TH is 200 mg/L as $CaCO_3$. No health-based guideline value is proposed by WHO for hardness in drinking-water. TH of all rivers were below 200 mg/L as $CaCO_3$.

Dissolved oxygen (DO) is one of the most important water quality parameters for assessment of surface water quality. An adequate amount of DO is essential for proper functioning of the aquatic ecosystem. The major sources of DO include reaeration from the atmosphere, photosynthetic oxygen production, and introduction of DO from other sources such as tributaries. Main causes of oxygen depletion are oxidation of organic material and other reduced matter in the water column, degassing of oxygen in supersaturated water, respiration by aquatic plants, and oxygen demand exerted by river bed sediments (Cox, 2003). Low concentration of DO can cause alterations in the kinds of aquatic organisms found in water bodies. Species that cannot tolerate low levels of DO will be replaced by a few kinds of pollution-tolerant organisms, such as worms and fly larvae. Nuisance algae and anaerobic organisms may also become abundant in waters with low levels of DO. DO of all rivers were observed above 6 mg/L except at one location of Brahmaputra River (SSR5). SSR5 is the confluence point of Brahmaputra and Bharalu River. Bharalu River is one of the most polluted rivers of Assam. It carries a large portion of the Guwahati city's municipal as well as other wastes and also serves as the natural drainage for storm water runoff.

Biochemical oxygen demand, (BOD) is one of the most significant and often used water quality parameters for assessing the pollution level of aquatic bodies (Basant et al., 2010). It is an approximate measure of bio-degradable organic matter present in water and defined by the amount of oxygen required for the aerobic micro-organisms present in the sample to oxidise the organic matter to a stable inorganic form (Sawyer and McCarty, 1978; Babu et al., 2006; Basant et al., 2010). Major sources of BOD include domestic, industrial and commercial wastewater, dead plants and animals, animal manure, leaves and woody debris and urban stormwater runoff (Jha et al., 2007). High BOD adversely affects the aquatic systems. It may cause water quality problems such as severe dissolved oxygen depletion and fish kills. BOD of all rivers was below 10 mg/L except Brahmaputra, Baralia, and Kolong River. Highest average BOD was observed in Brahmaputra River (17.12 mg/L).

Sodium (Na^+) is ubiquitous in water, due to the high solubility of its salts and the abundance of sodium-containing mineral deposits. Major anthropogenic sources of Na^+ in surface water include road salt, water treatment chemicals, domestic water softeners, and sewage effluents. Na^+ is not considered to be toxic but higher concentration may affect the taste of drinking water. No health based guideline has been provided for Na^+ by BIS and WHO. Average concentration of Na^+ in all rivers was below 10 mg/L.

Potassium (K^+) is an essential element in humans and generally occurs in rock and soil. It enters surface waters through weathering. Other sources include leachate from municipal landfill site and runoff from agricultural land. K^+ is an important constituent of synthetic fertilizers (NPK fertilizers) and limiting factor for the growth of terrestrial vegetation (Talling, 2010). It helps in nerve stimulus, muscle contractions, blood pressure regulation and protein dissolution. It is seldom, if ever, found in drinking water at levels that could be a concern for healthy humans. No health based guideline has been provided by BIS and WHO. Average concentration of K^+ in all rivers was below 5 mg/L.

Calcium ($Ca^{2+)}$ and magnesium (Mg^{2+}) occur naturally in surface water. These cations enter into stream due to weathering of rocks, and from the soil through seepage, leaching and runoff. Both Ca^{2+} and Mg^{2+} are essential to human health in several respects. Inadequate intake of either nutrient can result in adverse health consequences. Low Ca^{2+} levels are associated with increased risks of osteoporosis, nephrolithiasis (kidney stones), colorectal cancer, hypertension and stroke, coronary artery disease, insulin resistance and obesity. Inadequate intakes of Mg^{2+} may be associated with endothelial dysfunction, increased vascular reactions, elevated circulating levels of C-reactive protein and decreased insulin sensitivity. Low Mg^{2+} may also cause hypertension, coronary heart disease, type 2 diabetes mellitus and metabolic syndrome. As per BIS guidelines of drinking water, acceptable limit of Ca^{2+} is 75 mg/L and for Mg^{2+} acceptable limit is 30 mg/L. Average concentration of Ca^{2+} and Mg^{2+} in all rivers was well below the acceptable limits. Highest concentration of Mg^{2+} was observed in Manas River.

Fluoride (F^-) is a naturally occurring element found in the earth's crust. It occurs naturally in surface waters from the atmospheric deposition and weathering of fluoride-containing rocks and soils. Various anthropogenic activities such as chemical manufacturing plants and production of fluoridated chemicals also introduce fluoride in water. At low pH, fluoride forms complexes with metal ions and at higher pH, it exists as a single fluoride ion. Several epidemiological studies show that long-term exposure of fluoride causes adverse impacts on skeletal tissues (bones and teeth). Lower concentrations of F^- in drinking water provides protection against dental caries but higher concentration may give rise to mild dental fluorosis. As per BIS guidelines of drinking water, acceptable limit of F^- is 1 mg/L. As per WHO guidelines, concentrations of F^- should be ranged from 0.5 - 1 mg/L. Average concentration of F^- in all rivers was below 1 mg/L.

Chloride (Cl^-) is a naturally occurring major anion and generally present at low concentrations in natural surface waters. Major sources of Cl^- include rocks containing chlorides, effluent wastewater from wastewater treatment plants, wastewater from industries, agricultural runoff and road salting. Chloride is normally considered as a conservative ion because its movement is not retarded by the interaction of water with soils, sediments, and rocks. At low concentration Cl^- is non-toxic but higher concentration is noxious to Fish and aquatic bodies. Elevated concentration of Cl^- in drinking water causes the salty taste. As per BIS guidelines of drinking water, acceptable limit of Cl^- is 250 mg/L. No health-based guideline value has been proposed by WHO for Cl^- in drinking-water. Average concentration of Cl^- in all rivers was below 10 mg/L except in the Brahmaputra, Beki and Kameng River. Concentration of Cl^- was highest in Brahmaputra River. It was in the range of 23.84-733.22 with an average of 412.94 mg/L

Sulphate (SO_4^{2-}) is an important constituent of surface water. SO_4^{2-} may get into surface water from several sources including sewage treatment plants and industrial discharges such as tanneries, pulp mills, and textile mills, breakdown of leaves that fall into a stream, atmospheric deposition and rock or soil containing gypsum and other common minerals through which streams flow. Surface runoff from fertilized agricultural lands also contributes SO_4^{2-}. The excess amount of SO_4^{2-} in drinking water can have a laxative effect with Ca^{2+} and Mg^{2+}. No health-based guideline value has been derived by WHO for SO_4^{2-}. As per BIS guidelines of drinking water, acceptable limit of SO_4^{2-} is 200 mg/L. Average concentration of SO_4^{2-} in all rivers was below the permissible limit except in Brahmaputra River (18.06 mg/L - 474.06 mg/L).

Nitrate (NO_3^-) pollution of surface water is a major concern due to its various adverse impacts (Ribbe et al., 2008). High concentration of NO_3^- in drinking water may cause serious health issues such as blue baby syndrome and diarrhoea (Gupta et al., 2001; Ribbe et al., 2008). Major sources of nitrates include wastewater treatment plants, runoff from fertilized agriculture lands and animal manure storage areas. The impact of intensive agriculture on nitrate pollution has been reported in several studies (Ribbe et al., 2008). Nitrates from terrestrial sources transpire in rivers and streams more swiftly than other nutrients because they dissolve in water more readily. Excess concentration of nitrate negatively affects the surface water resources in the form of eutrophication, low dissolved oxygen levels and anoxic events. As per BIS guidelines of drinking water, acceptable limit of NO_3^- is 45 mg/L and as per WHO guidelines it is 50 mg/L. Average concentration of NO_3^- in all rivers was well below the permissible limits.

4.3 HEAVY METALS AND ASSOCIATED HEALTH RISK ASSESSMENT

Iron (Fe) is essential trace element required for good health. It helps in transportation of oxygen in the blood. High level of Fe can cause hemochromatosis which can harm liver, heart, pancreas. It can cause fatigue, weight loss, joint pain, stomach problem, and vomiting (Anderson, 1994). Loose intracellular iron can also promote DNA damage. Minimum daily requirement for iron vary person to person and it depends on age, sex, and physiological status. As per BIS guidelines of drinking water, acceptable limit of Fe is 0.3 mg/L. No health based guideline has been provided by WHO for Fe. Average concentration of Fe in Baralia, Pagladia and Beki River was below acceptable limit. Highest average concentration of Fe was observed in Kameng River (0.54mg/L).

Manganese (Mn) is necessary for humans to survive, as it is required for the functioning of many cellular enzymes. Its chronic exposure can cause manganese poisoning or Parkinson's disease (Goldhaber, 2003; Harmanescu et al., 2011). As per BIS guidelines of drinking water, acceptable limit of Mn is 0.1 mg/L. As per health based guidelines of WHO, permissible concentration of Mn is 0.4 mg/L. Average concentration of Mn in all rivers except Baralia River was below permissible limit. In Baralia River Mn ranged from 0.03 to 0.63 mg/L with an average value 0.15 mg/L.

Lead (Pb) is generally recognized as one of the most pervasive environmental health threat. It can damage mainly hematopoietic system, nervous system and renal system (Papanikolaou et al., 2005; Khan et al., 2013). It is known to be a potent inhibitor of heme synthesis. Pb compounds can also damage RBCs. It causes abdominal pain, vomiting, diarrhoea, collapse, high blood pressure and heart attack (Khan et al., 2013). As per WHO and BIS guidelines of drinking water, acceptable limit of Pb in drinking water is 0.01 mg/L. Average concentration of Pb in all rivers except Puthimari, was above permissible limit. Highest concentration (0.03 mg/L) of Pb was found in Baralia and Kolong River.

Copper (Cu) is an essential element for the function of many cellular enzymes (Tapiero et al., 2003). Long-term exposure to Cu can cause irritation of the nose, mouth and eyes and also causes headaches, stomach-aches, dizziness, vomiting, diarrhoea and can damage liver (Goldhaber, 2003; Harmanescu et al., 2011). As per BIS guidelines of drinking water, acceptable limit of Cu in drinking water is 0.05 mg/L. As per health based guidelines of WHO, permissible concentration of Cu is 2.0 mg/L. Average concentration of Cu in Kolong River was 0.07 mg/L and in Brahmaputra River, it was 0.06 mg/L. In all other rivers it was well below the permissible limit (0.05 mg/L).

Zinc (Zn) is necessary nutrient for body growth and development because it is a component of a wide variety of enzymes (Goldhaber, 2003). High levels of Zn can lead to stomach cramps, nausea, vomiting, skin irritations, gastrointestinal irritation and anaemia (Prasad, 1976; Fosmire, 1990). As per BIS guidelines of drinking water, acceptable limit of Zn is 5.0 mg/L. No health-based guideline value has been proposed by WHO for Zn in drinking water. Average Zn concentration in all rivers was well below the permissible limit.

Effect of chromium (Cr) on human health depends on its oxidation state. Cr (III) is an essential nutrient for human body but too much uptake can cause liver and kidney problems (Goldhaber, 2003). Cr (VI) is toxic to human health and may cause adverse health impact. Intravascular haemolysis and renal failure have also been reported from the high level Cr in the body (Lamson and Plaza, 2002). As per WHO and BIS guidelines of drinking water, acceptable limit of Cr is 0.05 mg/L. In Baralia and Kolong River, average concentration of Cr was above permissible limit. In Baralia River Cr was in the range of 0.02-0.11 mg/L and in Kolong River it was in the range of 0.00–0.29 mg/L.

Table 4.1(a). Statistical summary of physico-chemical parameters of Brahmaputra River

SS	pH	DO	TSS	TDS	EC	BOD	TH	TA	Na^+	K^+	Ca^{2+}	Mg^{2+}	F^-	Cl^-	SO_4^{2-}	NO_3^-
SSR1	7.9	9.1	117.5	43.0	0.2	6.8	81.0	77.0	3.8	2.0	31.1	23.8	0.1	23.8	18.9	0.1
SSR2	8.0	8.8	146.5	49.5	0.2	14.1	128.0	75.0	4.2	2.3	29.1	25.8	0.1	282.4	22.4	0.2
SSR3	8.0	8.6	114.0	121.0	0.2	20.1	103.0	77.0	4.1	2.2	26.7	25.7	0.2	109.8	22.1	0.3
SSR4	7.9	8.8	123.5	126.5	0.2	17.9	105.0	77.0	4.0	2.1	40.3	25.0	0.1	675.6	474.1	1.3
SSR5	7.2	2.3	142.0	150.5	0.6	29.7	133.0	177.0	34.7	14.3	85.4	32.2	0.4	733.2	18.1	1.3
SSR6	8.3	9.6	179.5	164.0	0.2	16.4	107.0	64.0	4.2	2.0	48.5	25.3	0.1	434.3	22.5	0.1
SSR7	8.2	8.9	113.0	138.0	0.2	18.2	140.0	78.0	4.4	2.0	49.2	25.0	0.1	240.2	21.8	0.4
SSR8	8.1	9.1	147.5	116.5	0.3	16.5	101.0	74.0	4.5	2.1	47.6	25.1	0.1	604.0	22.2	0.3
SSR9	8.1	9.0	116.5	36.0	0.2	14.6	140.0	85.0	4.5	2.1	47.1	24.9	0.2	613.1	22.2	0.8
Max	8.3	9.6	179.5	164.0	0.6	29.7	140.0	177.0	34.7	14.3	85.4	32.2	0.4	733.2	474.1	1.3
Min	7.2	2.3	113.0	36.0	0.2	6.8	81.0	64.0	3.8	2.0	26.7	23.8	0.1	23.8	18.1	0.1
Mean	8.0	8.2	133.3	105.0	0.3	17.1	115.3	87.1	7.6	3.4	45.0	25.9	0.2	412.9	71.6	0.5
SD	0.3	2.3	22.4	49.0	0.1	6.1	20.6	34.1	10.2	4.1	17.6	2.4	0.1	259.5	150.9	0.5
Skewness	-2.2	-2.9	1.1	-0.5	2.6	0.6	-0.2	2.9	3.0	3.0	1.6	2.7	2.4	-0.3	3.0	0.9
Kurtosis	5.8	8.6	0.9	-1.5	7.0	2.8	-1.1	8.4	9.0	9.0	3.5	7.6	6.0	-1.6	9.0	-0.7
CoV	0.0	0.3	0.2	0.5	0.5	0.4	0.2	0.4	1.3	1.2	0.4	0.1	0.5	0.6	2.1	0.9

Table 4.1(b). Statistical summary of physico-chemical parameters of Baralia River

SS	pH	DO	TSS	TDS	EC	BOD	TH	TA	Na^+	K^+	Ca^{2+}	Mg^{2+}	F^-	Cl^-	SO_4^{2-}	NO_3^-
SPBR1	7.2	6.9	132.0	112.0	0.2	12.9	68.0	92.0	4.9	1.2	14.5	14.3	0.2	1.0	15.0	0.0
SPBR2	7.5	7.2	40.0	192.0	0.2	15.6	68.0	92.0	4.9	0.9	7.0	16.0	0.0	2.0	16.5	0.1
SPBR3	7.5	7.2	60.0	160.0	0.3	21.0	60.0	92.0	4.4	1.2	6.8	15.5	0.2	0.5	26.5	0.2
SPBR4	7.9	6.9	130.0	184.0	0.2	18.0	70.0	82.0	4.6	1.2	13.2	19.1	0.3	2.0	20.7	0.5
SPBR5	7.2	6.9	388.0	120.0	0.2	14.1	70.0	88.0	4.4	1.1	11.1	15.8	0.0	1.0	15.3	0.4
SPBR6	7.5	6.8	214.0	154.0	0.2	9.6	68.0	92.0	4.8	1.1	14.2	16.9	0.0	2.5	19.9	0.6
SPBR7	7.5	7.8	264.0	155.0	0.2	9.3	69.0	93.0	5.8	1.0	15.2	16.6	0.0	3.5	20.9	0.7
SPBR8	7.5	8.8	275.0	156.0	0.2	7.5	70.0	94.0	6.8	1.0	16.2	16.5	0.1	4.5	21.9	0.9
SPBR9	7.6	6.2	326.0	444.0	0.2	8.4	60.0	92.0	4.9	1.1	14.1	15.9	0.2	1.5	14.8	1.0
Max	7.9	8.8	388.0	444.0	0.3	21.0	70.0	94.0	6.8	1.2	16.2	19.1	0.3	4.5	26.5	1.0
Min	7.2	6.2	40.0	112.0	0.2	7.5	60.0	82.0	4.4	0.9	6.8	14.3	0.0	0.5	14.8	0.0
Mean	7.5	7.2	203.2	186.3	0.2	12.9	67.0	90.8	5.0	1.1	12.5	16.3	0.1	2.1	19.0	0.5
SD	0.2	0.8	120.3	100.0	0.0	4.7	4.1	3.7	0.8	0.1	3.5	1.3	0.1	1.3	3.9	0.3
Skewness	0.5	1.4	0.1	2.6	1.7	0.5	-1.4	-2.1	1.8	-0.4	-1.0	1.1	0.4	0.9	0.6	-0.1
Kurtosis	1.5	2.5	-1.2	7.4	3.7	-0.9	0.4	4.5	3.0	-0.8	-0.4	2.8	-1.3	0.2	-0.2	-1.4
CoV	0.0	0.1	0.6	0.5	0.1	0.4	0.1	0.0	0.2	0.1	0.3	0.1	1.0	0.6	0.2	0.7

Table 4.1(c). Statistical summary of physico-chemical parameters of Puthimari River

SS	pH	DO	TSS	TDS	EC	BOD	TH	TA	Na^+	K^+	Ca^{2+}	Mg^{2+}	F^-	Cl^-	SO_4^{2-}	NO_3^-
SPPR1	7.6	6.7	76.0	206.0	0.7	2.4	156.0	168.0	4.4	4.1	22.7	30.3	0.3	1.5	28.8	0.2
SPPR2	7.8	7.6	118.0	244.0	0.2	9.6	64.0	62.0	2.5	0.7	12.8	15.9	0.3	1.5	38.8	0.4
SPPR3	7.6	7.1	118.0	156.0	0.2	6.9	66.0	86.0	3.0	1.0	16.0	14.1	0.0	2.0	24.4	0.1
SPPR4	7.4	7.2	170.0	184.0	0.2	7.2	70.0	82.0	4.6	1.2	13.2	13.7	0.3	2.0	27.4	0.8
SPPR5	7.8	7.1	148.0	148.0	0.2	7.8	86.0	80.0	3.4	1.1	14.1	14.5	0.1	1.0	28.3	0.5
SPPR6	7.6	7.1	178.0	126.0	0.2	3.3	72.0	86.0	2.5	1.1	13.4	17.5	0.0	1.5	32.1	0.4
Max	7.8	7.6	178.0	244.0	0.7	9.6	156.0	168.0	4.6	4.1	22.7	30.3	0.3	2.0	38.8	0.8
Min	7.4	6.7	76.0	126.0	0.2	2.4	64.0	62.0	2.5	0.7	12.8	13.7	0.0	1.0	24.4	0.1
Mean	7.6	7.1	134.7	177.3	0.3	6.2	85.7	94.0	3.4	1.5	15.4	17.7	0.2	1.6	30.0	0.4
SD	0.1	0.3	38.2	43.0	0.2	2.8	35.3	37.3	0.9	1.3	3.8	6.3	0.2	0.4	5.0	0.2
Skewness	-1.1	0.4	-0.4	0.6	2.4	-0.5	2.2	2.1	0.5	2.3	2.0	2.2	-0.2	-0.3	1.2	0.8
Kurtosis	1.4	2.0	-0.6	-0.5	6.0	-1.3	5.0	5.0	-1.9	5.6	4.1	5.0	-2.3	-0.1	1.8	0.8
CoV	0.0	0.0	0.3	0.2	0.7	0.4	0.4	0.4	0.3	0.8	0.2	0.4	0.9	0.2	0.2	0.6

Table 4.1(d). Statistical summary of physico-chemical parameters of Pagladia River

SS	pH	DO	TSS	TDS	EC	BOD	TH	TA	Na^+	K^+	Ca^{2+}	Mg^{2+}	F^-	Cl^-	SO_4^{2-}	NO_3^-
SSPaR1	7.6	8.8	36.0	80.0	0.3	3.9	80.0	50.0	3.7	1.8	32.2	12.4	0.1	3.4	1.2	0.7
SSPaR2	7.7	8.9	172.0	124.0	0.3	2.4	102.0	110.0	5.5	2.0	32.7	15.4	0.1	0.3	0.7	3.2
SSPaR3	7.6	8.8	150.0	90.0	0.3	9.6	94.0	111.0	5.6	2.0	30.6	15.2	0.1	10.4	8.2	0.3
SSPaR4	7.5	8.8	347.8	86.0	0.5	6.9	96.0	98.0	4.2	1.8	33.3	15.5	0.3	5.0	10.1	0.0
SSPaR5	7.5	8.7	216.0	180.0	0.4	7.2	96.0	120.0	5.7	2.3	41.0	14.5	0.1	1.5	10.0	0.6
SSPaR6	7.5	8.6	208.0	154.0	0.3	7.8	68.0	86.0	4.6	2.0	39.9	13.2	0.1	1.4	9.7	0.4
SSPaR7	7.5	8.8	200.0	160.0	0.3	3.3	72.0	74.0	4.3	2.0	36.5	13.0	0.1	1.3	6.9	0.3
SSPaR8	7.5	8.8	211.0	154.0	0.2	5.4	66.0	69.0	4.1	1.7	32.6	11.6	0.1	1.0	4.2	0.3
SSPaR9	7.5	8.9	210.0	137.0	0.2	3.3	60.0	46.0	4.0	1.4	30.6	10.9	0.1	0.9	4.4	0.1
Max	7.7	8.9	347.8	180.0	0.5	9.6	102.0	120.0	5.7	2.3	41.0	15.5	0.3	10.4	10.1	3.2
Min	7.5	8.6	36.0	80.0	0.2	2.4	60.0	46.0	3.7	1.4	30.6	10.9	0.1	0.3	0.7	0.0
Mean	7.5	8.8	194.5	129.4	0.3	5.5	81.6	84.9	4.6	1.9	34.4	13.5	0.1	2.8	6.2	0.7
SD	0.1	0.1	80.9	36.5	0.1	2.5	15.7	27.0	0.8	0.3	3.8	1.7	0.1	3.2	3.7	1.0
Skewness	1.8	-0.8	-0.1	-0.3	1.4	0.3	0.0	-0.2	0.5	-0.2	0.9	-0.2	2.3	2.0	-0.4	2.7
Kurtosis	3.5	0.8	2.9	-1.5	1.1	-1.2	-1.9	-1.4	-1.5	1.2	-0.6	-1.5	5.6	4.3	-1.5	7.8
CoV	0.0	0.0	0.4	0.3	0.3	0.4	0.2	0.3	0.2	0.1	0.1	0.1	0.5	1.1	0.6	1.5

Table 4.1(e). Statistical summary of physico-chemical parameters of Beki River

SS	pH	DO	TSS	TDS	EC	BOD	TH	TA	Na^{+}	K^{+}	Ca^{2+}	Mg^{2+}	F^{-}	Cl^{-}	SO_4^{2-}	NO_3^{-}
SSBeK1	7.6	9.9	132.0	112.0	0.2	0.6	115.0	92.0	2.4	2.1	22.4	14.3	0.1	6.0	7.4	0.1
SSBeK2	7.7	9.6	40.0	126.0	0.2	1.2	114.0	92.0	3.1	2.1	19.3	16.0	0.1	12.0	12.4	0.1
SSBeK3	7.6	9.7	60.0	160.0	0.3	2.8	137.0	92.0	3.0	2.0	29.4	15.5	0.2	14.0	11.0	0.1
SSBeK4	7.5	9.1	130.0	144.0	0.2	3.3	136.0	82.0	3.5	1.3	23.1	19.1	0.1	17.0	10.5	0.7
SSBeK5	7.5	8.3	388.0	200.0	0.3	11.1	140.0	88.0	5.7	1.9	20.2	15.8	0.1	22.0	11.6	0.9
SSBeK6	7.5	9.0	214.0	205.0	0.3	5.7	135.0	92.0	3.2	1.6	26.1	16.9	0.1	16.0	10.1	0.7
SSBeK7	7.5	10.2	264.0	195.0	0.3	4.7	130.0	90.0	3.2	1.6	25.0	16.6	0.1	17.0	7.5	0.1
SSBeK8	7.5	10.4	275.0	212.0	0.2	3.6	132.0	92.0	3.1	1.3	25.0	16.5	0.1	15.0	7.0	0.0
Max	7.7	10.4	388.0	212.0	0.3	11.1	140.0	92.0	5.7	2.1	29.4	19.1	0.2	22.0	12.4	0.9
Min	7.5	8.3	40.0	112.0	0.2	0.6	114.0	82.0	2.4	1.3	19.3	14.3	0.1	6.0	7.0	0.0
Mean	7.6	9.5	187.9	169.3	0.2	4.1	129.9	90.0	3.4	1.7	23.8	16.3	0.1	14.9	9.7	0.3
SD	0.1	0.7	118.8	38.9	0.0	3.3	10.0	3.5	1.0	0.3	3.3	1.4	0.0	4.6	2.1	0.4
Skewness	1.8	-0.6	0.4	-0.4	-0.5	1.5	-1.1	-2.1	2.2	-0.3	0.3	0.9	0.5	-0.7	-0.3	0.8
Kurtosis	3.2	-0.2	-0.7	-1.8	-1.5	2.9	-0.5	4.2	5.8	-1.8	-0.1	2.2	0.0	1.8	-1.8	-1.6
CoV	0.0	0.1	0.6	0.2	0.2	0.8	0.1	0.0	0.3	0.2	0.1	0.1	0.2	0.3	0.2	1.1

Table 4.1(f). Statistical summary of physico-chemical parameters of Manas River

SS	pH	DO	TSS	TDS	EC	BOD	TH	TA	Na^+	K^+	Ca^{2+}	Mg^{2+}	F^-	Cl^-	SO_4^{2-}	NO_3^-
SSMR1	7.4	7.1	55.0	205.0	0.4	3.0	182.2	130.0	4.5	2.3	25.0	29.2	0.1	8.3	17.4	1.5
SSMR2	7.5	7.2	60.0	235.0	0.3	4.9	169.0	109.0	4.7	3.3	20.0	29.0	0.2	13.1	18.7	0.3
SSMR3	7.6	7.2	110.0	210.0	0.4	4.4	150.1	104.0	4.9	2.4	12.4	29.1	0.1	14.1	19.2	2.4
SSMR4	7.6	7.1	85.0	235.0	0.4	5.4	158.2	127.0	3.9	3.0	15.4	29.2	0.1	4.6	19.0	0.5
SSMR5	7.5	7.3	575.0	210.0	0.3	12.8	120.4	62.0	3.8	7.1	5.7	25.9	0.3	11.1	36.4	1.9
SSMR6	7.4	7.5	730.0	185.0	0.3	9.3	133.7	63.0	3.6	7.3	11.0	25.9	0.4	6.2	36.3	0.9
SSMR7	7.9	7.3	170.0	285.0	0.4	3.2	165.0	131.0	3.3	4.4	16.8	30.0	0.1	4.7	22.1	0.6
SSMR8	7.7	7.5	175.0	250.0	0.4	3.5	177.6	147.0	3.4	7.7	21.6	30.1	0.1	11.6	22.6	1.4
SSMR9	7.9	7.4	185.0	280.0	0.4	3.5	182.2	150.0	4.2	7.3	24.6	30.2	0.1	11.6	23.8	2.0
Max	7.9	7.5	730.0	285.0	0.4	12.8	182.2	150.0	4.9	7.7	25.0	30.2	0.4	14.1	36.4	2.4
Min	7.4	7.1	55.0	185.0	0.3	3.0	120.4	62.0	3.3	2.3	5.7	25.9	0.1	4.6	17.4	0.3
Mean	7.6	7.3	238.3	232.8	0.4	5.5	159.8	113.7	4.0	5.0	16.9	28.7	0.1	9.5	24.0	1.3
SD	0.2	0.2	242.9	34.2	0.1	3.3	21.7	32.7	0.6	2.4	6.5	1.7	0.1	3.6	7.3	0.7
Skewness	0.5	-0.2	1.6	0.4	-0.6	1.7	-0.8	-0.8	0.3	0.1	-0.4	-1.3	1.5	-0.3	1.3	0.1
Kurtosis	-1.0	-1.5	1.2	-0.9	-0.6	2.0	-0.3	-0.6	-1.5	-2.3	-0.7	0.2	1.4	-1.6	0.2	-1.4
CoV	0.0	0.0	1.0	0.2	0.2	0.6	0.1	0.3	0.2	0.5	0.4	0.1	0.7	0.4	0.3	0.6

Table 4.1(g). Statistical summary of physico-chemical parameters of Kolong River

SS	pH	DO	TSS	TDS	EC	BOD	TH	TA	Na^+	K^+	Ca^{2+}	Mg^{2+}	F^-	Cl^-	SO_4^{2-}	NO_3^-
SPKR1	7.0	6.9	146.0	100.0	0.2	9.6	26.0	58.0	5.0	1.1	1.7	9.5	0.0	3.5	0.3	0.2
SPKR2	7.1	6.2	128.0	184.0	0.2	12.9	32.0	54.0	4.8	1.5	1.5	13.7	0.0	3.0	3.6	0.3
SPKR3	7.6	6.0	136.0	142.0	0.2	15.9	40.0	60.0	5.3	1.5	1.1	14.2	0.3	4.5	5.5	0.2
SPKR4	6.9	5.9	114.0	132.0	0.2	13.2	44.0	62.0	5.5	1.4	1.6	14.7	0.2	7.5	3.6	0.4
SPKR5	6.9	6.1	112.0	114.0	0.2	11.7	44.0	60.0	5.7	1.2	1.7	15.0	0.1	9.5	3.2	0.4
SPKR6	7.0	5.8	113.0	113.0	0.2	10.5	40.0	54.0	6.0	1.2	1.8	15.0	0.1	7.0	3.6	0.2
SPKR7	7.0	5.7	122.0	108.0	0.2	10.9	42.0	46.0	6.2	1.5	1.8	15.7	0.1	6.5	4.0	0.1
SPKR8	6.9	6.0	114.0	124.0	0.2	11.4	42.0	42.0	7.0	1.7	3.1	15.2	0.1	5.0	0.6	0.1
SPKR9	6.8	6.2	119.0	126.0	0.2	7.5	34.0	36.0	4.6	1.5	2.6	10.1	0.0	3.5	0.0	0.0
Max	7.6	6.9	146.0	184.0	0.2	15.9	44.0	62.0	7.0	1.7	3.1	15.7	0.3	9.5	5.5	0.4
Min	6.8	5.7	112.0	100.0	0.2	7.5	26.0	36.0	4.6	1.1	1.1	9.5	0.0	3.0	0.0	0.0
Mean	7.0	6.1	122.7	127.0	0.2	11.5	38.2	52.4	5.6	1.4	1.9	13.7	0.1	5.6	2.7	0.2
SD	0.2	0.3	11.8	24.9	0.0	2.4	6.2	9.1	0.8	0.2	0.6	2.3	0.1	2.2	1.9	0.1
Skewness	2.4	1.8	1.1	1.6	0.6	0.2	-1.1	-0.8	0.7	-0.6	1.2	-1.4	1.5	0.6	-0.4	0.1
Kurtosis	6.4	4.3	0.4	3.3	-0.5	0.9	0.3	-0.6	0.1	-0.6	1.4	0.3	1.4	-0.7	-1.2	-1.1
CoV	0.0	0.1	0.1	0.2	0.1	0.2	0.2	0.2	0.1	0.1	0.3	0.2	1.0	0.4	0.7	0.6

Table 4.1(h). Statistical summary of physico-chemical parameters of Kameng River

SS	pH	DO	TSS	TDS	EC	BOD	TH	TA	Na^+	K^+	Ca^{2+}	Mg^{2+}	F^-	Cl^-	SO_4^{2-}	NO_3^-
SSKaR1	7.6	9.3	116.0	210.0	0.1	6.8	60.7	22.0	2.5	2.4	11.6	7.8	0.1	14.6	5.3	2.3
SSKaR2	7.9	9.3	130.0	195.0	0.1	2.9	60.9	20.0	2.5	3.2	11.5	7.9	0.1	10.7	5.4	2.0
SSKaR3	8.0	9.2	140.0	70.0	0.1	5.1	57.7	19.0	2.4	3.1	10.3	7.8	0.1	11.9	5.3	1.4
SSKaR4	6.8	8.3	105.0	335.0	0.1	3.8	89.0	30.0	3.4	2.4	12.8	13.9	0.1	13.9	7.0	2.1
SSKaR5	7.1	8.1	50.0	295.0	0.1	6.3	97.4	32.0	3.2	2.3	14.6	14.8	0.2	13.2	8.3	0.7
SSKaR6	7.1	8.2	75.0	605.0	0.1	11.1	59.5	28.0	3.2	2.5	2.3	13.1	0.1	10.6	6.8	1.7
SSKaR7	7.3	8.3	195.0	370.0	0.1	7.8	58.2	27.0	2.8	3.4	5.3	11.0	0.1	13.6	7.0	6.5
SSKaR8	7.3	8.3	260.0	190.0	0.1	3.5	89.4	26.0	2.6	1.5	18.9	10.3	0.1	10.3	7.8	3.6
SSKaR9	7.4	8.2	220.0	115.0	0.1	6.2	97.4	29.0	2.7	3.6	4.9	9.4	0.3	7.0	8.5	1.0
Max	8.0	9.3	260.0	605.0	0.1	11.1	97.4	32.0	3.4	3.6	18.9	14.8	0.3	14.6	8.5	6.5
Min	6.8	8.1	50.0	70.0	0.1	2.9	57.7	19.0	2.4	1.5	2.3	7.8	0.1	7.0	5.3	0.7
Mean	7.4	8.6	143.4	265.0	0.1	5.9	74.5	25.9	2.8	2.7	10.2	10.7	0.1	11.7	6.8	2.4
SD	0.4	0.5	68.9	160.7	0.0	2.6	18.1	4.6	0.4	0.7	5.2	2.7	0.1	2.4	1.3	1.8
Skewness	0.2	0.8	0.5	1.1	-0.5	0.9	0.4	-0.4	0.5	-0.5	0.0	0.4	2.5	-0.9	-0.2	1.9
Kurtosis	-0.7	-1.7	-0.7	1.7	-0.8	1.0	-2.3	-1.2	-1.4	0.0	-0.6	-1.5	6.2	0.7	-1.6	4.1
CoV	0.1	0.1	0.5	0.6	0.1	0.4	0.2	0.2	0.1	0.2	0.5	0.3	0.6	0.2	0.2	0.7

Table 4.2 (a). Statistical summary of heavy metals of Brahmaputra River

SS	Fe	Mn	Pb	Cu	Zn	Cr
SSR1	0.03	0.00	0.02	0.03	0.04	0.05
SSR2	0.03	0.00	0.02	0.04	0.04	0.06
SSR3	0.41	0.00	0.01	0.09	0.16	0.06
SSR4	1.00	0.00	0.03	0.02	0.04	0.01
SSR5	0.93	0.16	0.03	0.19	0.34	0.05
SSR6	0.51	0.00	0.01	0.10	0.19	0.06
SSR7	0.08	0.00	0.03	0.04	0.06	0.06
SSR8	0.04	0.00	0.02	0.04	0.04	0.06
SSR9	0.08	0.00	0.01	0.03	0.25	0.04
Max	1.00	0.16	0.03	0.19	0.34	0.06
Min	0.03	0.00	0.01	0.02	0.04	0.01
Mean	0.35	0.02	0.02	0.06	0.13	0.05
SD	0.39	0.05	0.01	0.05	0.11	0.02
Skewness	0.94	3.00	0.00	1.81	0.95	-2.11
Kurtosis	-0.75	9.00	-1.71	3.23	-0.33	4.71
CoV	1.14	3.00	0.43	0.85	0.87	0.33

Table 4.2 (b). Statistical summary of heavy metals of Baralia River

SS	Fe	Mn	Pb	Cu	Zn	Cr
SPBR1	0.03	0.09	0.00	0.02	0.03	0.02
SPBR2	0.08	0.09	0.00	0.01	0.01	0.06
SPBR3	0.11	0.02	0.01	0.01	0.03	0.08
SPBR4	0.18	0.18	0.07	0.02	0.02	0.05
SPBR5	0.63	0.28	0.07	0.06	0.04	0.05
SPBR6	0.22	0.17	0.03	0.02	0.02	0.09
SPBR7	0.2	0.17	0.06	0.02	0.01	0.08
SPBR8	0.19	0.17	0.03	0.02	0.02	0.1
SPBR9	0.34	0.19	0.02	0.02	0.01	0.11
Max	0.63	0.28	0.07	0.06	0.04	0.11
Min	0.03	0.02	0.00	0.01	0.01	0.02
Mean	0.22	0.15	0.03	0.02	0.02	0.07
SD	0.18	0.07	0.03	0.01	0.01	0.03
Skewness	1.70	-0.18	0.36	2.48	0.55	-0.44
Kurtosis	3.53	0.60	-1.56	7.01	-0.55	-0.35
CoV	0.81	0.49	0.87	0.67	0.50	0.40

Table 4.1(c). Statistical summary of heavy metals of Puthimari River

SS	Fe	Mn	Pb	Cu	Zn	Cr
SPPR1	0.46	0.00	0.00	0.05	0.03	0.00
SPPR2	0.22	0.00	0.00	0.05	0.06	0.01
SPPR3	0.18	0.01	0.00	0.04	0.05	0.01
SPPR4	0.09	0.00	0.00	0.05	0.04	0.01
SPPR5	0.99	0.01	0.00	0.04	0.05	0.01
SPPR6	0.82	0.02	0.01	0.07	0.04	0.01
Max	0.99	0.02	0.01	0.07	0.06	0.01
Min	0.09	0.00	0.00	0.04	0.03	0.00
Mean	0.46	0.01	0.00	0.05	0.05	0.01
SD	0.37	0.01	0.00	0.01	0.01	0.00
Skewness	0.65	0.86	2.45	1.37	0.00	-2.45
Kurtosis	-1.61	-0.30	6.00	2.50	-0.25	6.00
CoV	0.80	1.22	2.45	0.22	0.23	0.49

Table 4.2 (d). Statistical summary of heavy metals of Pagladia River

SS	Fe	Mn	Pb	Cu	Zn	Cr
SSPaR1	0.08	0.00	0.01	0.00	0.00	0.00
SSPaR2	0.13	0.04	0.01	0.00	0.00	0.00
SSPaR3	0.16	0.03	0.01	0.00	0.00	0.02
SSPaR4	0.20	0.06	0.01	0.00	0.06	0.00
SSPaR5	0.31	0.05	0.01	0.00	1.64	0.00
SSPaR6	0.26	0.03	0.02	0.00	0.00	0.00
SSPaR7	0.23	0.04	0.03	0.00	0.00	0.00
SSPaR8	0.19	0.00	0.03	0.00	0.00	0.00
SSPaR9	0.25	0.00	0.01	0.05	0.03	0.00
Max	0.31	0.06	0.03	0.05	1.64	0.02
Min	0.08	0.00	0.01	0.00	0.00	0.00
Mean	0.20	0.03	0.02	0.01	0.19	0.00
SD	0.07	0.02	0.01	0.02	0.54	0.01
Skewness	-0.26	-0.22	1.19	3.00	2.99	3.00
Kurtosis	-0.28	-1.44	-0.45	9.00	8.96	9.00
CoV	0.35	0.82	0.57	3.00	2.83	3.00

Table 4.2 (e). Statistical summary of heavy metals of Beki River

SS	Fe	Mn	Pb	Cu	Zn	Cr
SSBeK1	0.03	0.01	0.01	0.02	0.03	0.02
SSBeK2	0.20	0.01	0.00	0.01	0.05	0.01
SSBeK3	0.13	0.00	0.01	0.01	0.06	0.07
SSBeK4	0.15	0.02	0.00	0.02	0.04	0.03
SSBeK5	0.36	0.06	0.01	0.06	0.04	0.10
SSBeK6	0.22	0.05	0.01	0.03	0.02	0.01
SSBeK7	0.19	0.03	0.01	0.02	0.02	0.01
SSBeK8	0.17	0.02	0.01	0.02	0.01	0.01
Max	0.36	0.06	0.01	0.06	0.06	0.10
Min	0.03	0.00	0.00	0.01	0.01	0.01
Mean	0.18	0.03	0.01	0.02	0.03	0.03
SD	0.09	0.02	0.00	0.02	0.02	0.03
Skewness	0.51	0.77	-1.44	1.98	0.17	1.50
Kurtosis	2.24	-0.45	0.00	4.57	-0.91	1.12
CoV	0.51	0.83	0.62	0.67	0.50	1.05

Table 4.2 (f). Statistical summary of heavy metals of Manas River

SS	Fe	Mn	Pb	Cu	Zn	Cr
SSMR1	0.05	0.00	0.00	0.05	0.03	0.00
SSMR2	0.22	0.00	0.00	0.05	0.06	0.01
SSMR3	0.18	0.01	0.00	0.04	0.05	0.01
SSMR4	0.09	0.00	0.00	0.05	0.04	0.01
SSMR5	0.99	0.01	0.00	0.04	0.05	0.01
SSMR6	0.82	0.02	0.01	0.07	0.04	0.01
SSMR7	0.17	0.00	0.00	0.04	0.04	0.00
SSMR8	0.31	0.01	0.00	0.04	0.06	0.00
SSMR9	0.51	0.04	0.03	0.03	0.02	0.04
Max	0.99	0.04	0.03	0.07	0.06	0.04
Min	0.05	0.00	0.00	0.03	0.02	0.00
Mean	0.37	0.01	0.00	0.05	0.04	0.01
SD	0.33	0.01	0.01	0.01	0.01	0.01
Skewness	1.11	1.67	2.51	1.16	-0.37	2.10
Kurtosis	-0.06	2.95	6.34	2.33	-0.31	5.43
CoV	0.90	1.32	2.28	0.25	0.31	1.22

Table 4.2 (g). Statistical summary of heavy metals of Kolong River

SS	Fe	Mn	Pb	Cu	Zn	Cr
SPKR1	0.47	0.1	0.02	0.02	0.01	0.02
SPKR2	0.27	0.21	0.03	0.06	0.02	0.19
SPKR3	0.26	0.24	0.04	0.09	0.02	0.29
SPKR4	0.47	0.24	0.03	0.06	0.02	0.19
SPKR5	0.24	0.23	0.03	0.06	0.02	0.17
SPKR6	0.26	0.22	0.03	0.06	0.02	0.17
SPKR7	0.28	0.26	0.03	0.07	0.02	0.21
SPKR8	0.34	0.24	0.03	0.03	0.01	0.04
SPKR9	0.27	0.01	0.03	0.14	0.00	0.00
Max	0.47	0.26	0.04	0.14	0.02	0.29
Min	0.24	0.01	0.02	0.02	0.00	0.00
Mean	0.32	0.19	0.03	0.07	0.02	0.14
SD	0.09	0.08	0.01	0.03	0.01	0.10
Skewness	1.29	-1.80	0.00	1.11	-1.50	-0.33
Kurtosis	0.00	2.47	4.00	2.28	1.47	-1.06
CoV	0.28	0.43	0.17	0.53	0.47	0.69

Table 4.2 (h). Statistical summary of heavy metals of Kameng River

SS	Fe	Mn	Pb	Cu	Zn	Cr
SSKaR1	0.14	0.00	0.03	0.00	0.00	0.00
SSKaR2	0.29	0.00	0.01	0.00	0.00	0.00
SSKaR3	1.13	0.00	0.00	0.00	0.00	0.00
SSKaR4	0.54	0.00	0.02	0.01	0.01	0.02
SSKaR5	1.34	0.04	0.02	0.01	0.00	0.02
SSKaR6	0.45	0.03	0.03	0.01	0.00	0.01
SSKaR7	0.27	0.00	0.03	0.01	0.00	0.01
SSKaR8	0.56	0.05	0.04	0.00	0.00	0.00
SSKaR9	0.17	0.00	0.03	0.00	0.00	0.00
Max	1.34	0.05	0.04	0.01	0.01	0.02
Min	0.14	0.00	0.00	0.00	0.00	0.00
Mean	0.54	0.01	0.02	0.00	0.00	0.01
SD	0.42	0.02	0.01	0.01	0.00	0.01
Skewness	1.18	1.08	-0.82	0.27	3.00	0.82
Kurtosis	0.27	-0.75	0.35	-2.57	9.00	-1.08
CoV	0.78	1.55	0.52	1.19	3.00	1.30

4.4 CHAPTER CONCLUSION

This chapter presented the 1st phase of the present study. In this phase, based on a survey of study area 68 sampling sites in Brahmaputra River and its seven tributaries namely Baralia, Puthimari, Pagladia, Beki, Manas, Kolong and Kameng River were identified. Water samples were collected from these sampling locations and analysed in the laboratory. Analysis showed that the average concentration of most of the physico-chemical parameters was well below the permissible limits laid down by BIS and WHO, except Baralia, Kolong and Brahmaputra rivers. General characteristics of physico-chemical parameters and health risk associated with heavy metals were also discussed in brief.

The observed data must be processed and presented in a manner that helps in the understanding of the water quality variability, taking into consideration the natural processes and characteristics of a water body, and that allows the impact of human activities to be understood and the consequences of management action to be predicted. This is not to say that water quality information must be presented in a way which requires the user to appreciate the full complexity of aquatic systems. The information should provide the user with the understanding necessary to meet the objectives behind the monitoring programme. In the next phases of the study, various multivariate statistical techniques and water quality indices have been applied for evaluation of data sets with a view to getting better information about water quality and design of monitoring network for effective management of water resources.

5

Chapter

Identification of latent pollution sources

This chapter presents the results of

- Cluster analysis
- Discriminant analysis
- Principal component analysis
- Concluding remarks

5.1 CLUSTER ANALYSIS (CA)

CA was performed on the water quality dataset to statistically categorize the monitoring stations into various homogeneous and distinct groups/clusters, based on the similarities among the sampling sites. Hierarchical agglomerative clustering, which is one of the most common method, was used for classifying the monitoring locations. Hierarchical cluster analysis (HCA) resulted in the extraction of a Dendrogram, which aids in the classification through pictorial representation. (Jung et al., 2016; Kim et al., 2007). Classification of groups was accomplished based on the observed water quality datasets for Brahmaputra River and its tributaries separately as shown in Fig. 5.1 (a) and Fig. 5.1 (b) respectively.

The sampling locations of Brahmaputra River were classified into three clusters; first cluster comprising of all sampling sites except SSR4 and SSR5. SSR4 was included in second cluster, whereas SSR5 was in third cluster. Sampling sites SSR4 and SSR5 were located near the confluence point of Brahmaputra and Bharalu River, therefore representing sites of medium (MP) and high pollution (HP) as Bharalu river is one of the most polluted rivers of Assam and adversely affecting the water quality of Brahmaputra River. First cluster, thus represents sites having relatively less pollution (LP).

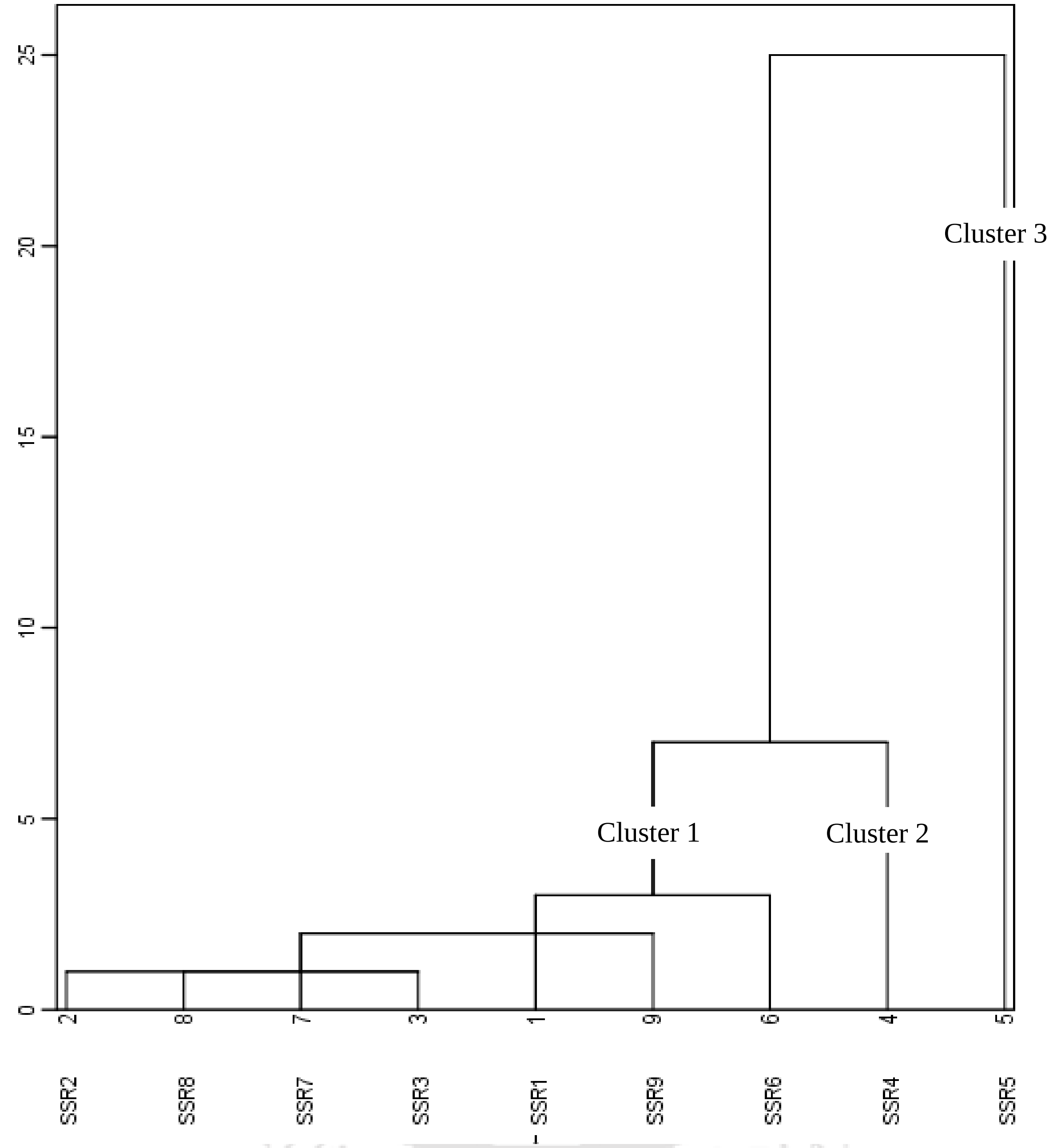


Fig. 5.1(a). Dendrogram for Brahmaputra River

Dendrogram for tributaries of Brahmaputra River; Baralia, Puthimari, Pagladia, Beki, Manas, Kolong and Kameng River has been shown in Fig. 5.1 (b). Cluster 1 included all sampling sites of Beki, Pagladia and Kameng River, representing sites of relatively low pollution (LP). All sampling sites of Manas River and first sampling site of Puthimari River (SPPR1) were included in second cluster and was thus categorized as sites of medium pollution (MP). Similarly, third cluster corresponded to sites of Baralia River and Puthimari River (except SPPR1), classified as high pollution (HP) sites.

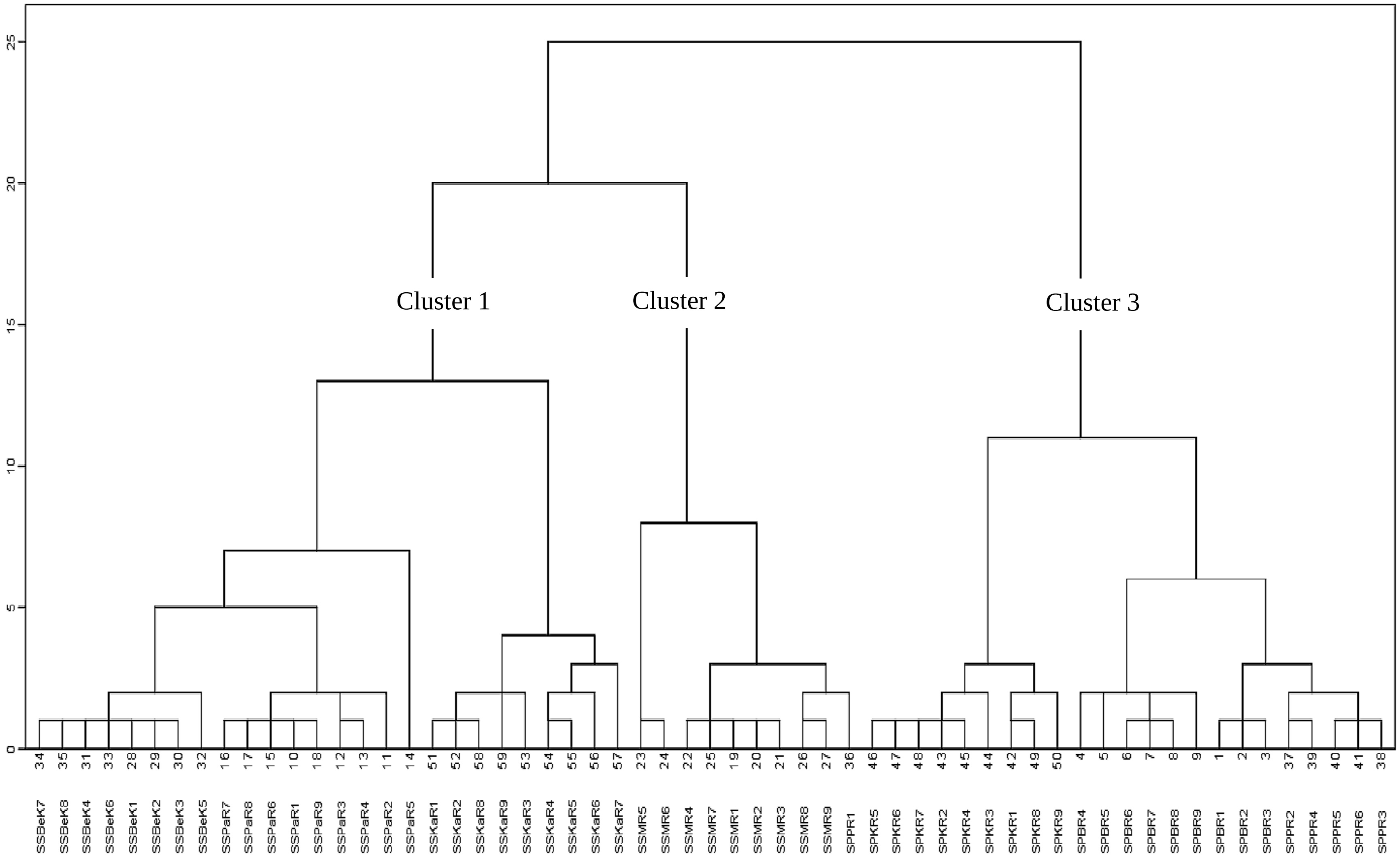


Fig. 5.1(b). Dendrogram for tributaries of Brahmaputra River

The results of HCA therefore suggest a cost-effective approach for sampling by providing reliable categorization of the monitoring locations. This will further help in deciding an optimum future sampling strategy of the surface waters.

5.2 DISCRIMINANT ANALYSIS (DA)

DA was employed on the raw dataset after classifying the sites into three classes (from CA). DA can be performed in two modes: Standard mode and Stepwise mode. Standard mode takes into account all the predictive variables for the construction of discriminating functions (DFs). The stepwise method starts with a model that doesn't include any of the predictors. At each step, the predictor with the largest F to enter value that exceeds the entry criteria is added to the model. Stepwise method is helpful when number of predictive variable is large. Since the total number of predictive variable (water quality parameter) is 22 hence stepwise method was applied to choose best variables in the model.

The measured values of Wilk's lambda and chi-square statistic of discriminant function obtained from stepwise mode shown in Table 5.1 (a) for Brahmaputra River and in Table 5.1 (b) for its tributaries. Wilks' Lambda is the ratio of within-groups sums of squares to the total sums of squares. This is the proportion of the total variance in the discriminant scores not explained by differences among groups. It represents the discriminating ability of the function for the separation of cases into groups. A lambda of 1.00 occurs when observed group means are equal (all the variance is explained by factors other than difference between those means), while a small lambda occurs when within-groups variability is small compared to the total variability. A small lambda indicates that group means appear to differ. Here, the Lambda of has a significant value (Sig. = 0,000) (Table 5.1); thus, the group means are to differ. Chi-square statistic is used to test the hypothesis that the means of the functions listed are equal between the considered groups.

Table 5.1 (a). Wilk's lambda and chi-square test for the spatial discriminant analysis of water quality for Brahmaputra River

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 2	.000	98.441	10	.000
2	.000	34.636	4	.000

Table 5.1 (b). Wilk's lambda and chi-square test for the spatial discriminant analysis of water quality for tributaries

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 2	.005	283.389	16	.000
2	.085	129.713	7	.000

Classification function coefficients obtained from stepwise mode of DA are shown in Table 5.2. Five water quality parameters (Na^{+} Ca^{2+}, Mg^{2+}, $SO4^{2-}$, and Pb) for Brahmaputra River (Table 5.2(a)) and eight parameters (DO, TA, K^{+} Ca^{2+}, Mg^{2+}, Cl^{-}, SO_4^{2-}, and Mn) for its tributaries (Table 5.2(b)), were identified as the "best" predictive variables (discriminant variables) that explain the variability of water quality in three cluster. Classification functions, using five discriminant variables yielded the corresponding classification matrix assigning 78% of the cases correctly for Brahmaputra River and for its tributaries eight discriminant parameters yielded the corresponding classification matrix assigning 100% of the cases correctly. Therefore, DA allowed a reduction in the dimensionality of the large data set, delineating a few indicator parameters responsible for large variations in the water quality.

Table 5.2 (a). Classification Function Coefficients for Brahmaputra River

	Cluster		
	1.00	2.00	3.00
Na^{+}	6666.854	-128384.156	23056.347
Ca^{2+}	247.004	-4669.402	764.442
Mg^{2+}	9304.110	-175924.211	29543.248
SO_4^{2-}	-6046.194	115253.160	-19467.884
Pb	-105170.988	2019368.654	-338699.044
(Constant)	-69053.399	-24799744.502	-727579.603

Table 5.2 (b). Classification Function Coefficients for tributaries

	Cluster		
	1.00	2.00	3.00
DO	35.706	30.283	28.278
TA	-0.581	-0.717	-0.349
K^+	7.821	14.051	4.288
Ca^{2+}	0.271	-0.311	-0.634
Mg^{2+}	7.566	14.289	6.446
Cl^-	-1.192	-2.379	-1.839
SO_4^{2+}	-1.848	-2.227	-1.000
Mn	-43.147	-145.453	26.761
(Constant)	-190.457	-267.713	-123.287

Table 5.3 (a). Classification Results (Brahmaputra River)

			Predicted Group Membership			
		Cluster	1.00	2.00	3.00	Total
Original	Count	1.00	7	0	0	7
		2.00	0	1	0	1
		3.00	0	0	1	1
	%	1.00	100.0	.0	.0	100.0
		2.00	.0	100.0	.0	100.0
		3.00	.0	.0	100.0	100.0
Cross-validated	Count	1.00	7	0	0	7
		2.00	1	0	0	1
		3.00	1	0	0	1
	%	1.00	100.0	.0	.0	100.0
		2.00	100.0	.0	.0	100.0
		3.00	100.0	.0	.0	100.0

Table 5.3 (b). Classification Results (tributaries)

			Predicted Group Membership			
		Cluster	1.00	2.00	3.00	Total
Original	Count	1.00	26	0	0	26
		2.00	0	10	0	10
		3.00	0	0	23	23
	%	1.00	100.0	.0	.0	100.0
		2.00	.0	100.0	.0	100.0
		3.00	.0	.0	100.0	100.0
Cross-validated	Count	1.00	26	0	0	26
		2.00	0	10	0	10
		3.00	0	0	23	23
	%	1.00	100.0	.0	.0	100.0
		2.00	.0	100.0	.0	100.0
		3.00	.0	.0	100.0	100.0

5.3 PRINCIPAL COMPONENT ANALYSIS (PCA)

PCA provides information on the most significant parameters to describe the entire data set with a minimum loss of original information (Iscen et al., 2008). In present study, PCA was performed on the normalized data sets (zero mean and unit variance) separately for the three different groups for tributaries of Brahmaputra River (Cluster 1, Cluster 2, and Cluster 3) as described by CA. PCA could not be applied on the water quality dataset for Brahmaputra River due to insufficient data.

For extracting valuable factor, eigenvalue-one criterion or Kaiser criterion (Kaiser 1960; Kim and Mueller 1987, Wang et al., 2013) was used. An Eigenvalue determines the importance of the component. Components with the maximum Eigenvalues are the most momentous. Varimax rotation was applied to maximize the variance of each factor and for better interpretation of results.

Analysis showed that for first cluster six principal components had eigenvalues greater than one. These six principal components explain 77.92 % of the total variance in the data sets. First component accounted for 23.85 % of the total variance. Second component contributed 16.39 % of the total variance. Third component explained 12.92 % of the total variance. Fourth, fifth and sixth components explained 10.61 %, 7.13% and 7.01 % of the total variance, respectively. Five principal components were obtained with eigenvalues >1 summing 91.68 % of the total variance in the data set of second cluster. First component was associated with 32.92 % of the total variance, second component was associated with 17.83 % of the total variance and third component was associated with 17.25 % of the total variance. Last two components explained approximately 23.68 % of the total variance. For the data set representing the third cluster, PCA yielded five principal components explaining 79.29 % of the total variance. First component contributed 33.81 % of the total variance. Second and third component contributed 16.15 % and 10.96 % of the total variance respectively. Fourth and fifth components explained nearly 18.37 % of the total variance. The Eigenvalues of all components, percentage variance explained by each component, cumulative percentage for three clusters are shown in Table 5.4 (a) – 5.4 (c). Scree plot of the eigenvalues of principal components are shown in Fig. 5.2 (a) – 5.2 (c). Scree plot is another useful method normally used by researchers to identify the number of principal components (up to and including the first one after the brake) to be retained.

Table 5.5 (a) shows the rotated component matrix for water quality data sets of cluster 1. First component has strong positive loading on TA, Ca^{2+}, EC and Mg^{2+} and moderate positive loading on TH. This factor is associated with mineral component of river water (Vega et al., 1998). Second component has strong positive loading on Cl^-, Cr and TH, moderate positive loading on Cu and SO_4^{2-}. This component may be associated with ‘‘soil leaching’’ processes. Third component has strong negative loading on pH, strong positive loading on TDS and moderate negative loading on DO. The strong negative loading in pH and moderate negative loading in DO is due to anaerobic conditions in the river from the loading of high dissolved organic matter, which results in formation of ammonia and organic acids leading to a decrease in pH (Shrestha and Kazama, 2007). Fourth component has strong positive loading on TSS and moderately positive loadings on BOD, Na^+ and Mn. Strong positive loading on TSS may be associated with erosion of rocks from upland areas, erosion of banks and surface runoff. The positive correlation with BOD indicates the loading of partially decayed organic matters from forest and agriculture areas (Shrestha and Kazama, 2007). Fifth component is strongly loaded with Zn and sixth component is strongly loaded with F^-.

Table 5.4 (a). Total Variance Explained (Cluster 1)

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	6.442	29.283	29.283	6.442	29.283	29.283	5.248	23.854	23.854
2	3.560	16.182	45.465	3.560	16.182	45.465	3.605	16.387	40.241
3	3.034	13.793	59.258	3.034	13.793	59.258	2.843	12.921	53.161
4	1.625	7.389	66.646	1.625	7.389	66.646	2.333	10.607	63.768
5	1.288	5.854	72.500	1.288	5.854	72.500	1.569	7.133	70.901
6	1.192	5.418	77.917	1.192	5.418	77.917	1.544	7.016	77.917

Table 5.4 (b). Total Variance Explained (Cluster 2)

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	8.487	38.578	38.578	8.487	38.578	38.578	7.243	32.921	32.921
2	5.082	23.101	61.680	5.082	23.101	61.680	3.922	17.827	50.748
3	3.230	14.684	76.364	3.230	14.684	76.364	3.795	17.250	67.998
4	2.179	9.906	86.269	2.179	9.906	86.269	2.768	12.582	80.581
5	1.191	5.414	91.683	1.191	5.414	91.683	2.443	11.103	91.683

Table 5.4 (c). Total Variance Explained (Cluster 3)

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	8.682	39.462	39.462	8.682	39.462	39.462	7.437	33.805	33.805
2	3.371	15.325	54.787	3.371	15.325	54.787	3.552	16.145	49.950
3	2.296	10.436	65.223	2.296	10.436	65.223	2.412	10.963	60.913
4	1.735	7.884	73.107	1.735	7.884	73.107	2.184	9.926	70.840
5	1.360	6.182	79.289	1.360	6.182	79.289	1.859	8.450	79.289

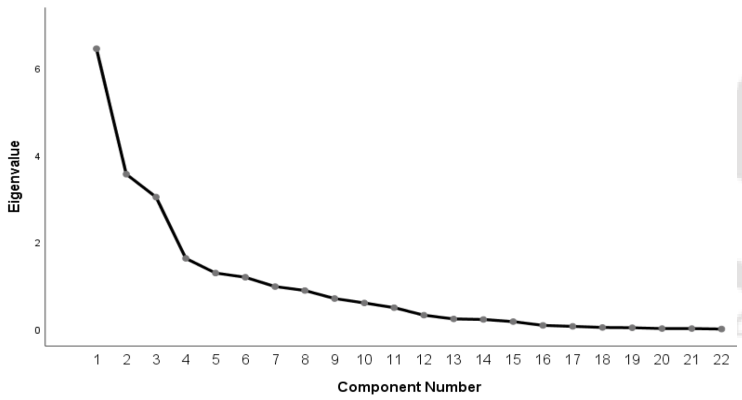


Fig. 5.2 (a) Scree Plot (Cluster 1)

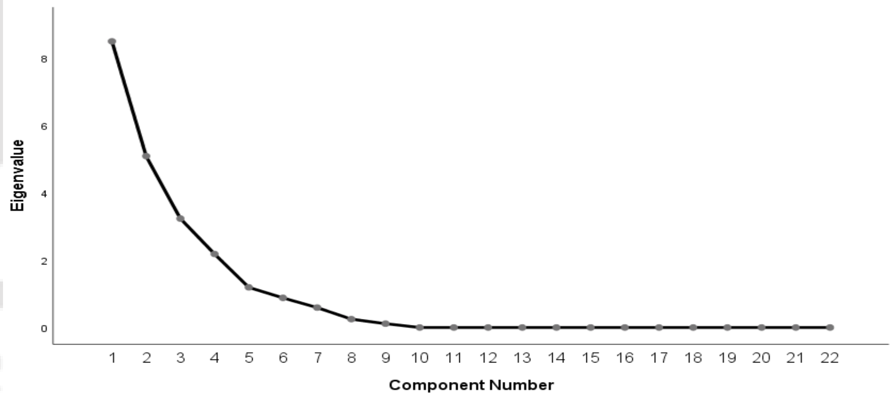


Fig. 5.2 (b) Scree Plot (Cluster 2)

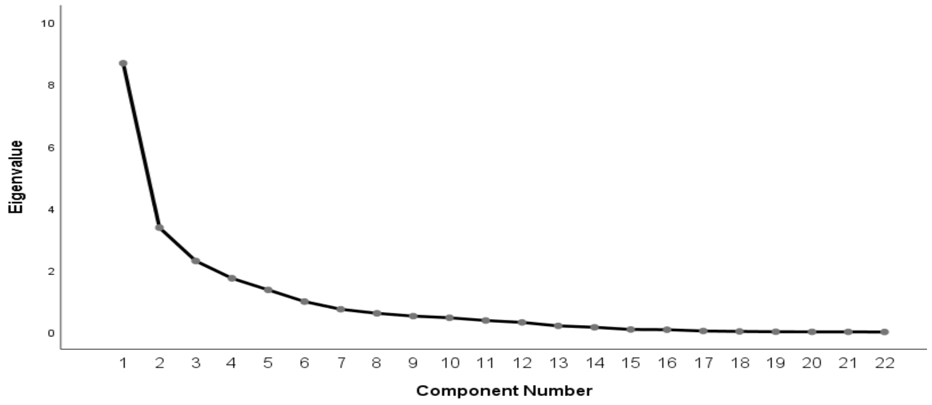


Fig. 5.2 (c) Scree Plot (Cluster 3)

Rotated component matrix for 2nd cluster is shown in Table 5.5 (b). First principal component has strong positive loadings on TSS, BOD, Fe and SO_4^{2-} and strong negative loadings on TH, Ca^{2+}, Mg^{2+} and TA. This factor is also moderately loaded with K^+ (positive loading) and TDS (negative loading). This factors represents the ‘‘physicochemical’’ source of the variability. Second component has strong positive loading on Pb, Mn and Cr. This component is associated with metal group (Singh et al., 2005). Third component is strongly loaded with EC and Cl^- and moderately loaded with DO and NO_3^-. This component represents influences from nonpoint sources such as agricultural runoff. Fourth component is strongly loaded with Cu and Zn, and fifth component is strongly loaded with Na.

Rotated component matrix for the water quality data set of 3rd cluster is shown in Table 5.5 (c). First component has strong positive loading on TH, TA, Ca^{2+}, SO_4^{2-}, EC, pH, and DO. This factor can be thus interpreted as a mineral component of the river water (Vega et al., 1998, Singh et al., 2004). Second component has strong positive loading on Mn and Pb, and moderate positive loading on Cr and Na. This factor represents the metal group (Singh et al., 2005). Third component is associated with NO_3^- and TSS. This factor represents the non-point source of pollution. Non-point sources of nitrate in this region is mainly runoff from fertilized agriculture lands and animal manure storage areas. Fourth component has strong positive loading on Fe. Fe in ground water is a major problem of Assam. Application of Fe contaminated water in various activities and discharge of that wastewater in streams may be the source of Fe.

Table 5.5 (a) Rotated Component Matrix (Cluster 1)

	Component					
	1	2	3	4	5	6
TA	**.850**	.179	-.285	.143	.198	.042
Ca	**.842**	-.277	-.264	.173	.142	-.087
EC	**.822**	-.047	-.246	.329	.203	.157
K	**-.801**	-.179	-.027	.012	.274	.301
Mg	**.769**	.522	.146	-.092	.086	.052
Nit	-.574	-.156	.297	.187	-.088	-.233
Fe	-.452	.118	.281	-.101	.402	-.002
Cl	-.294	**.870**	.091	-.078	-.105	-.027
Cr	.070	**.801**	.004	.189	.030	.019
TH	.533	**.720**	-.111	-.095	-.102	.218
Cu	.128	.685	-.033	.264	-.196	-.322
Sulf	.196	.632	.087	.071	.242	.408
pH	-.010	-.100	**-.928**	.017	.022	-.054
TDS	-.310	.180	**.780**	-.032	.096	-.258
Pb	-.296	-.406	.634	.255	-.294	.042
DO	.286	.219	-.591	-.448	-.242	-.205
TSS	.211	.147	-.113	**.805**	-.221	.055
BOD	-.205	.162	.341	.656	.332	.074
Na	.546	-.072	-.021	.586	.406	-.114
Mn	.477	.205	.243	.533	.239	.098
Zn	.253	-.119	-.052	.114	**.747**	-.023
F	-.008	.035	-.037	.100	-.065	**.948**

Table 5.5 (b) Rotated Component Matrix (Cluster 2)

	Component				
	1	2	3	4	5
TH	-.929	.223	-.066	.107	.076
Sulf	.927	.190	-.211	-.069	.202
Fe	.914	.326	-.029	.023	.142
BOD	.870	-.033	.351	-.184	.111
TSS	.844	.203	.157	-.279	.357
Mg	-.819	.020	-.333	.443	-.063
Ca	-.791	.297	-.376	.038	-.058
F	.749	.169	-.515	-.267	-.031
TA	-.701	.107	-.548	.393	-.051
Pb	-.029	.977	.035	.097	.136
Mn	.186	.913	.226	.167	.150
Cr	.019	.864	.262	.191	-.059
Zn	.141	-.650	.492	.045	.148
EC	-.225	-.024	-.856	.325	-.223
Cl	-.038	.161	.855	.220	-.196
DO	.174	.325	.661	.031	.626
Nit	.159	.358	.658	.339	-.251
Cu	.255	-.174	-.239	-.897	.063
pH	-.274	.242	-.120	.806	.396
TDS	-.502	.192	.048	.599	.492
Na	-.247	.120	.098	-.091	-.924
K	.502	.449	.031	.222	.600

Table 5.5 (c) Rotated Component Matrix (Cluster 3)

	Component				
	1	2	3	4	5
TH	**.896**	-.090	.135	.245	.020
TA	**.893**	-.006	.214	-.100	-.024
Ca	**.832**	-.197	.393	.137	-.053
Sulf	**.820**	-.366	.106	.290	.209
EC	**.814**	-.012	-.201	-.306	.016
pH	**.769**	-.138	.166	.229	.395
DO	**.738**	-.239	.225	-.065	-.199
K	-.693	.374	-.239	-.017	-.048
Mg	.667	.493	.017	.081	.258
Cu	-.659	-.035	-.058	.475	.028
Cl	-.641	.419	-.046	-.147	.007
Mn	-.306	**.895**	.073	-.138	.027
Pb	-.047	**.844**	.094	.053	-.130
Cr	-.438	.608	-.063	-.138	.419
Na	-.355	.587	.053	-.577	-.174
Nit	.368	.219	**.811**	-.008	.188
TSS	.284	.385	**.723**	.192	-.211
BOD	.126	.401	-.619	-.366	.334
TDS	.214	-.156	.607	-.230	.512
Fe	-.084	.053	.178	**.802**	-.184
Zn	.450	-.351	-.130	.604	.258
F	-.001	-.013	-.013	-.002	.877

5.4 CHAPTER CONCLUSION

In this 2nd phase of the study, different MSTs were used to evaluate the spatial variations in the surface water quality of Brahmaputra River and its tributaries. The important conclusions from the study were drawn as follows:

- Agglomerative hierarchical CA grouped 16 monitoring sites into three groups (low pollution, moderate pollution, and high pollution) based on their similarity of water quality characteristics.
- The results of CA can be used to reduce the need for numerous sampling sites and to optimize water quality monitoring program design with lower costs.
- DA, as an important data reduction method identified five water quality parameters (Na^+ Ca^{2+}, Mg^{2+}, $SO4^{2-}$, and Pb) for Brahmaputra River and eight parameters (DO, TA, K^+ Ca^{2+}, Mg^{2+}, Cl^-, SO_4^{2-}, and Mn) for its tributaries as the most important discriminating variables responsible for spatial variations.
- Although the PCA did not result in considerable data reduction, it helped extract and identify the factors responsible for variations in river water quality at different sampling sites.

This study showed the feasibility and reliability of the combined use of these multivariate statistical techniques in river water quality research.

6

Chapter

Water quality evaluation employing information entropy

This chapter covers spatial variation of water quality in terms of

- Entropy weighted WQI (EWQI)
- Entropy-weighted Heavy Metal Contamination Index (EHCI)
- Entropy weighted irrigation water quality index (EIWQI)
- TOPSIS
- Concluding Remarks

6.1 SPATIAL VARIATION OF WATER QUALITY IN TERMS OF 'EWQI'

Surface water quality monitoring programs evaluate a broad array of physical, chemical and biological water quality parameters including heavy metals. This necessitates the integration of these large and complex datasets into meaningful results that can represent the overall water quality status of a waterbody and can also be presented to planners and decision makers to take remedial action during an event of pollution. This led to the evolution of Water Quality Indices (WQIs) which aggregate a large set of measured parameters into a single numeric value (Zandbergen and Hall 1998). Crucial steps involved in the development of WQIs are the selection of parameters, weighing factors reflecting the importance of each parameter, and the final aggregation into a numerical score by establishing a rating scale for each parameter (Abbasi and Abbasi, 2012). Selection and weighing factors of parameters however, depend on the particular end use of the index which may be for drinking, recreation or irrigation purposes (Sharma et al., 2006; Misaghi et al., 2017). Many indices have been formulated and developed but there is no globally accepted index. The determination of weights by Delphi method, analytical hierarchy process (AHP) and expert survey method, which have been in use since the last few decades are tedious and time consuming exercise and often lead to deviation of weights due to subjective factors (Li et al., 2011; Wu et al., 2015). Water quality parameters being random variables, their probability distribution affects the probability distribution of the index and therefore its utility (Landwehr, 1979). Under such circumstances, computation of fixed weights of indices on the basis of inherent information of the indices could be capable of eliminating subjective disturbances (Li et al., 2011). Such information can be described by information entropy. Entropy weighted WQIs (EWQIs) are an improvement over conventional WQIs which are otherwise based on assignment of weights to parameters on the basis of personal judgements and expert opinion (Amiri et al., 2014; Fagbote et al., 2014).

In the present study, EWQI has been used for evaluation of water quality along the stretch of rivers with respect to drinking water quality standards. Spatial variation of water quality of Brahmaputra River and its tributaries are shown in Fig. 6.1 (a)-(h). However, before leaping on to the results of the EWQI, it is of paramount importance to develop an insight on the relationship between information entropy value and entropy weight with the physico-chemical parameters included in the study. Generally, parameters with low entropy value corresponding to higher entropy weights have the maximum influence on the overall water quality (Gorgij et al., 2017; Islam et al., 2017). The entropy weights of each physicochemical parameter for all rivers has been depicted in the Table 6.1. Spatial variation of water quality in terms of EWQI has been shown in Fig. 6.1 (a)-(h).

Table 6.1 Weights of physicochemical parameter

River	**pH**	**DO**	**TDS**	**EC**	**BOD_5**	**TH**	**TA**	**Na^+**	**K^+**	**Ca^{+2}**	**Mg^{+2}**	**F^-**	**Cl^-**	**SO_4^{2-}**	**NO_3^-**
Baralia	0.07	0.06	0.07	0.05	0.08	0.03	0.05	0.09	0.05	0.07	0.05	0.12	0.07	0.07	0.07
Kolong	0.11	0.07	0.08	0.07	0.04	0.04	0.04	0.07	0.05	0.06	0.05	0.12	0.08	0.07	0.06
Puthimari	0.03	0.03	0.04	0.18	0.04	0.11	0.06	0.06	0.09	0.09	0.10	0.06	0.03	0.04	0.05
Kameng	0.04	0.11	0.06	0.06	0.06	0.11	0.05	0.06	0.03	0.04	0.08	0.14	0.03	0.07	0.08
Manas	0.00	0.02	0.03	0.05	0.07	0.10	0.02	0.04	0.05	0.06	0.08	0.10	0.11	0.13	0.14
Beki	0.08	0.05	0.22	0.04	0.07	0.08	0.02	0.05	0.04	0.04	0.03	0.06	0.06	0.05	0.11
Pagladia	0.10	0.03	0.06	0.09	0.06	0.05	0.05	0.05	0.03	0.09	0.04	0.08	0.10	0.05	0.12
Brahmaputra	0.01	0.01	0.03	0.13	0.02	0.02	0.06	0.16	0.18	0.04	0.04	0.06	0.02	0.18	0.05

The grading of water quality has been done in accordance with the scale proposed by Jian-Hua et al., (2011). At all sampling sites, except SPBR3, SPBR4, SSBeK5 and SSR5, water quality ranged from "excellent" (EWQI < 50) to "good" (EWQI between 50 and 100). Water quality at sampling sites SPBR3 and SPBR4 of Baralia River, SSBeK5 of Beki River and SSR5 of Brahmaputra River were found to be in "poor" category (EWQI>100). Sampling sites SPBR3 and SPBR4 were located near densely populated market area of Rangia town receiving untreated domestic wastewater through open drains, due to lack of well-connected drainage system. Water of Baralia River is also used for washing clothes, bathing of pets, and fishing which also contribute to the degradation of river water quality. Another important factor contributing to pollution is the dumping of municipal solid waste (MSW) along the street sides and near the banks of the river. At numerous locations along the river banks, MSW was found to be dumped either as thin non-contiguous layers detached from the water or as thick contiguous heaps in contact with the flowing water. Thus, leachates from the wastes were directly or indirectly affecting the water quality one way or the other. Sampling site SSBeK5 was located in the proximity of Barpeta which is one of the most densely populated districts in the state of Assam. Although, the Beki River flows through the outskirts of the Barpeta city, considerably high BOD and Cl^- was recorded at the fifth location (SSBeK 5) suggesting possible contamination from nearby domestic and agricultural activities. Sampling point SSR5 is located at Bharalumukh, the converging point of Bharalu and Brahmaputra River. Bharalu River flows through densely populated residential, industrial and commercial areas of Guwahati city and meets the Brahmaputra river at Bharalumukh. Rapid urbanization of its catchment area in recent years has deteriorated it to a large extent due to unabated encroachment and dumping of garbage into it. Also, Bharalu receives both domestic and industrial wastes directly through several drains discharging all its pollutants to Brahmaputra River, thus polluting it to a great extent.

It was further observed that water quality at sampling sites in the upper reaches of the rivers where the population density was sparse and human activities were minimal (mostly agricultural activities were observed). Water quality at these locations were found to be either in "excellent" or "good" category. Furthermore, due to the effect of dilution and self-purification processes of rivers, the EWQI started decreasing in the downstream sections of the rivers (Bu et al., 2010).

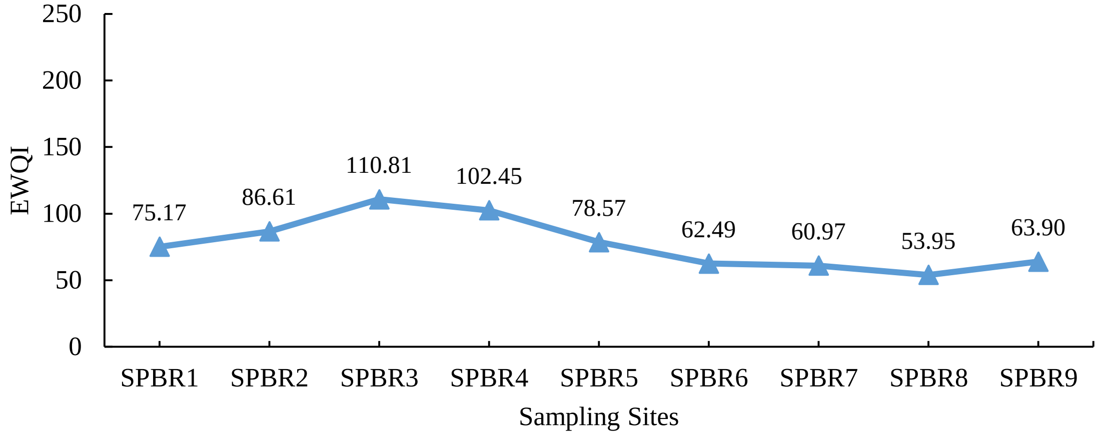


Fig. 6.1 (a). Spatial variation of water quality of Baralia River

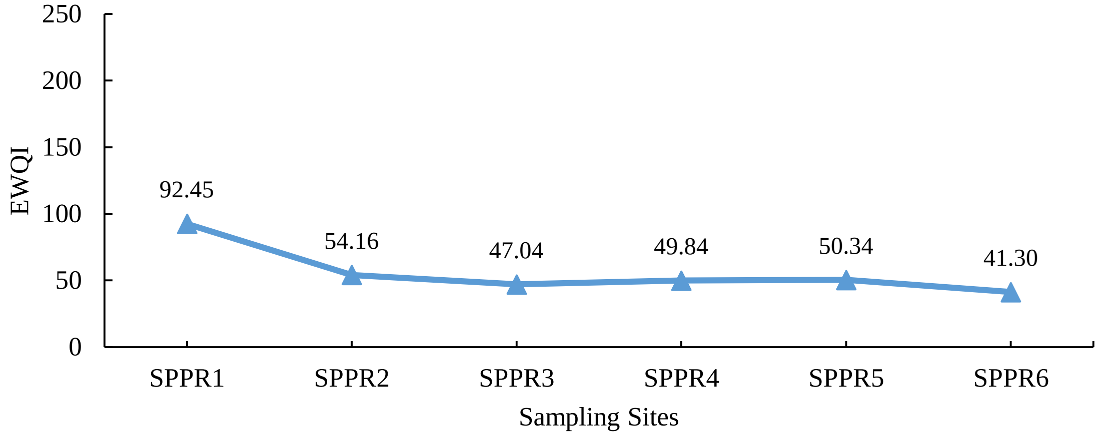


Fig. 6.1 (b). Spatial variation of water quality of Puthimari River

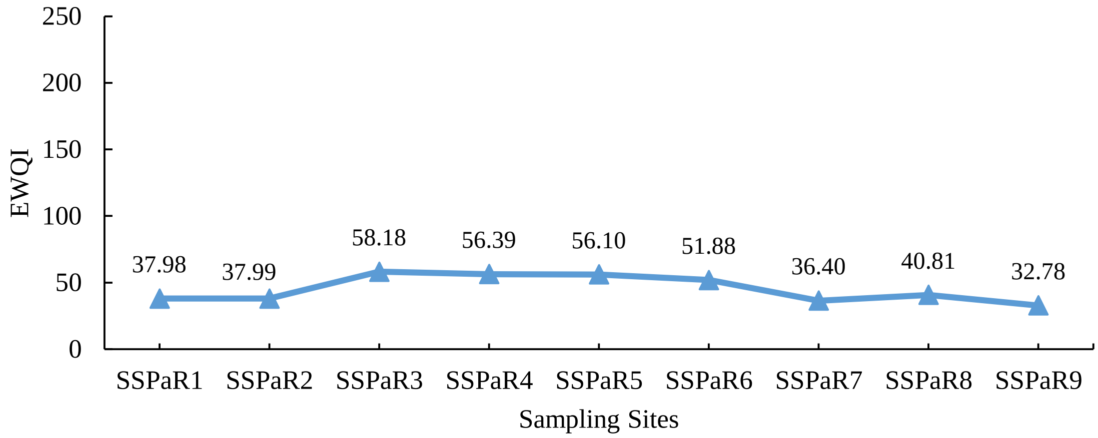


Fig. 6.1 (c). Spatial variation of water quality of Pagladia River

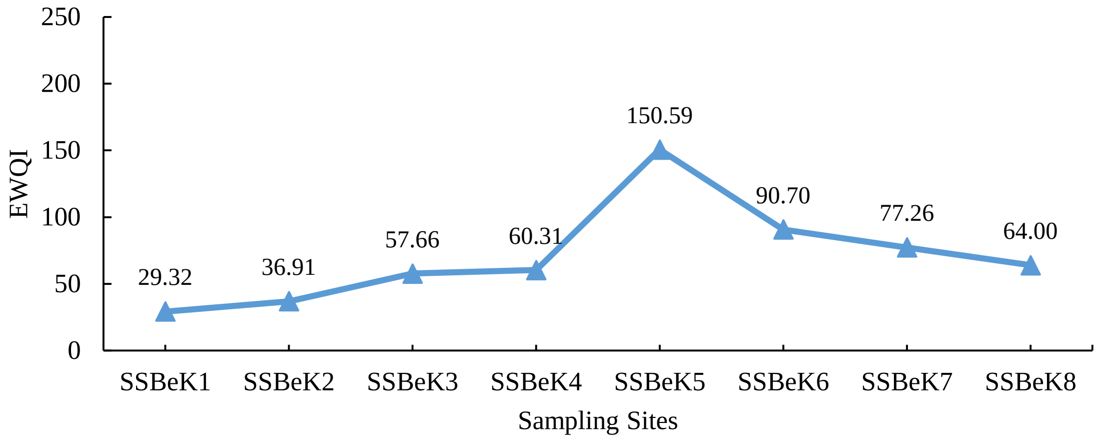


Fig. 6.1 (d). Spatial variation of water quality of Beki River

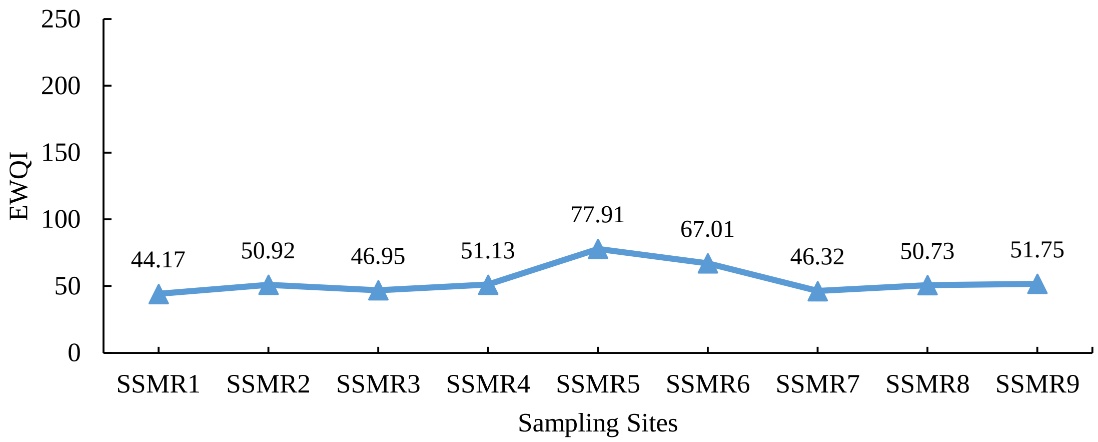


Fig. 6.1 (e). Spatial variation of water quality of Manas River

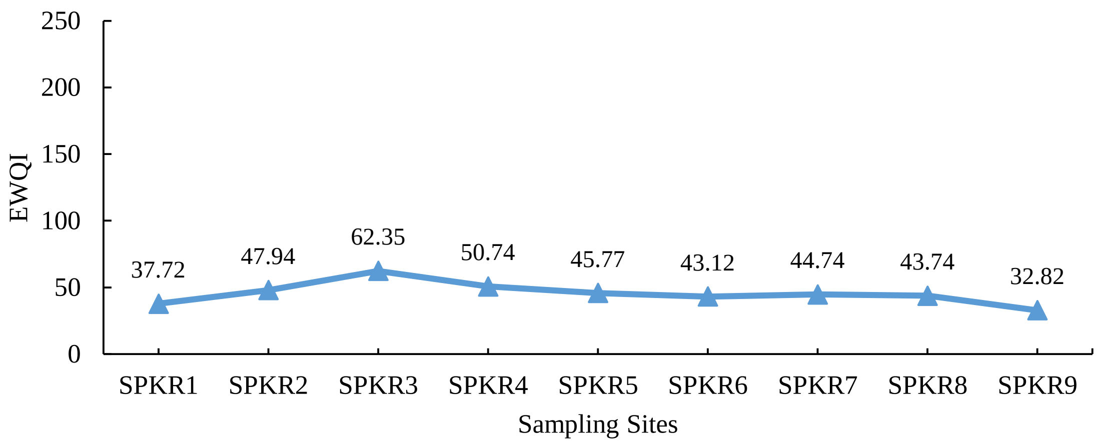


Fig. 6.1 (f). Spatial variation of water quality of Kolong River

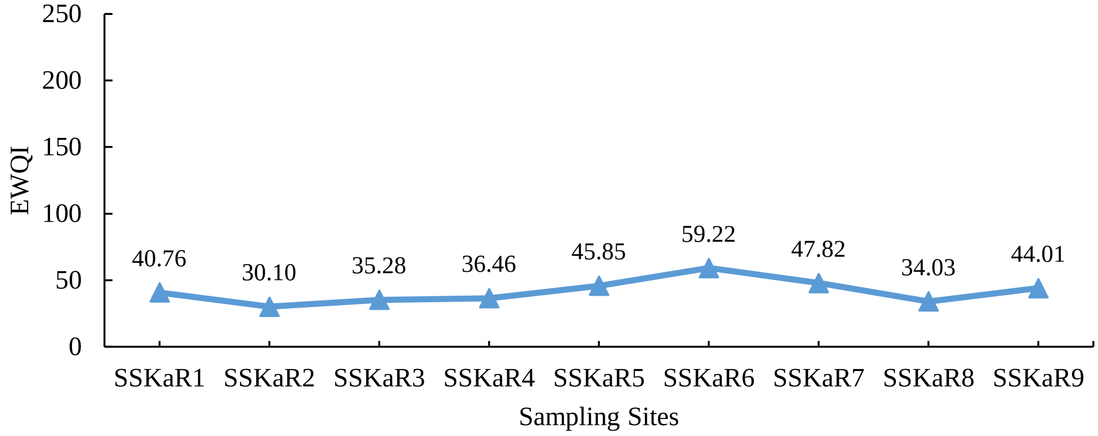


Fig. 6.1 (g). Spatial variation of water quality of Kameng River

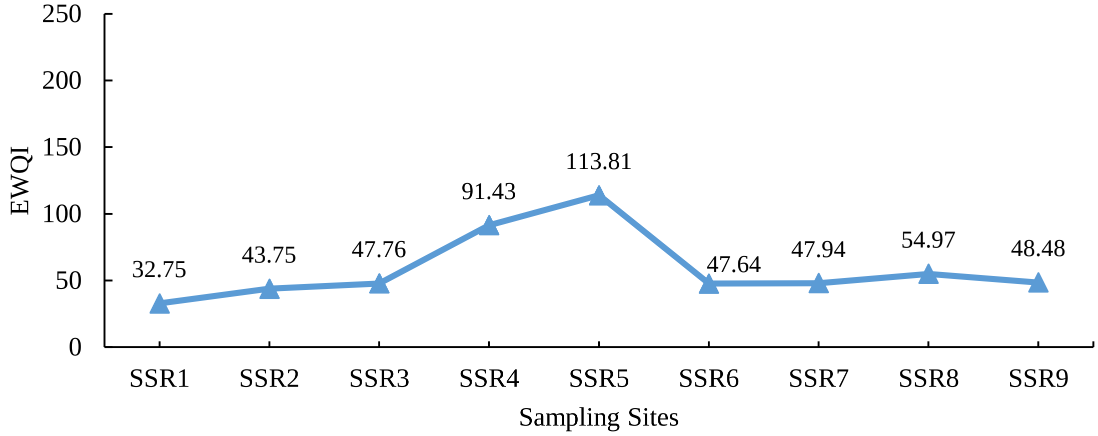


Fig. 6.1 (h). Spatial variation of water quality of Brahmaputra River

6.2 HEAVY METALS AND ITS INDEXING

Heavy metal contamination has become a global concern attributing to their abundance in aquatic bodies arising from either natural or anthropogenic activities. Quantification of heavy metal contamination has been done earlier by researchers employing Heavy Metal Pollution Index (HPI), Contamination Index (CI), Heavy Metal Evaluation Index (HEI) and Modified-Heavy Metal Pollution Index (m-HPI).

The HPI which revolves around the weighted arithmetic mean method of indexing is given as per the following equation, Eq. (6.1):

$$\mathrm{HPI} = \frac{\sum_{i=1}^{n} W_i Q_i}{\sum_{i=1}^{n} W_i} \tag{6.1}$$

Where, the weightage of each evaluated heavy metal is taken as the value inversely proportional to its maximum recommended standard with a proportionality constant of 1 ($\frac{1}{S_i}$). The sub-indexing (Q_i) has been done as per the following equation:

$$Q_i = \sum_{i=1}^{n} \frac{(M_i(-)I_i)}{(S_i - I_i)} \times 100 \quad (6.2)$$

Where, M_i denoted the monitored value of the i^{th} heavy metal; I_i denoted the ideal value of the i^{th} heavy metal and S_i denoted the upper permissible standard value of the i^{th} heavy metal.

As per the Eq. (6.2), the HPI suffers from two major drawbacks. The numerator which only takes into account the numerical difference between the monitored value and the desirable concentration of the heavy metal disqualifies concentrations of heavy metals very lower than their desirable concentrations thus maximizing their influence on the water quality. Therefore, for a metal such as Fe having its concentration 0.1 mg/L at one location and 0.5 mg/L at another would have the same influence on the overall water quality of that location with its desirable concentration being 0.3 mg/L as per the Eq. (6.2). Such a result becomes trivial in the perspective of a researcher. Secondly, the upper permissible limit in the absence of any external source is not specified in IS: 10500 (2012) for most metals such as Fe, Ag, Cd, Ni, etc. This also arises chaos in evaluating the denominator of the sub-index (Q_i).

The evaluation of the CI was developed at the Geological Survey of Slovak Republic by Bodis and Rapant and was further refined at the Geological Survey of Finland (Backman et al., 1998). This index was applied for assessment of heavy metal pollution in ground water in Finland and Slovakia by Backman et al., (1998) which is given as per the following equation (Eq. (6.3)):

$$CI = \sum_{i=1}^{n} C_{fi} \quad (6.3)$$

Where,

$$C_{fi} = \frac{C_{ai}}{C_{ni}} - 1 \quad (6.4)$$

And C_{fi} represents the contamination factor for the parameter and C_{ai} represents the analysed value of the i^{th} parameter. C_{ni} is the upper permissible limit of the i^{th} monitored parameter.

The contamination index discards contamination factors having negative values thus making it inappropriate for spatio-temporal assessments of water quality locations with parameters having concentrations lower than their upper permissible limits. Although, the contamination index clearly identifies the suitability of a source for drinking purpose in the absence of any alternate source, yet is ineffective in distinguishing between desirable concentration and upper permissible concentration.

In another study conducted by Edet and Offiong (2002), CI and HPI were applied at Akpabuyo-Odukpani area, Lower Cross River Basin (southeastern Nigeria). The two indices gave extreme classification results although they depicted good correlation with the heavy metal concentrations. Subsequently, a heavy metal evaluation index (HEI) was proposed as per the following equation:

$$\mathrm{HEI} = \sum_{i=1}^{n} {}^{H_c}/_{H_{mac}} \tag{6.5}$$

Where, H_c denoted the monitored value of i^{th} heavy metal and H_{mac} denoted the maximum admissible concentration of the heavy metal. The HEI developed to overcome deficiency in the conflicts between HPI and CI, classified water quality with respect to multiple of the mean as criteria. Furthermore, the HPI and CI scales were modified with respect to their multiples of mean. Such modifications deemed it possible to compare the three indices, however, the HEI was adopted due to its simplicity.

This method lacks assignment of weightages to heavy metals and completely ignores the plausible errors that can be caused from different quantity grades of quality. Consequently, the range of the results would be unnaturally large. The highlight in the HEI involves the classification of water quality and modification of the existing scales of HPI and CI by multiples of mean approach. The classification scale of this method is highly incongruous as the water quality is classified relative to the index's mean values. In such circumstances, if the heavy metal concentrations are very high in a water body, it will be classified as less and medium or less, medium and highly contaminated by the multiple of mean approach.

Recently, a modified HPI developed by Chaturvedi et al., (2018) tried to overcome the shortcomings of the HPI by introducing a modified sub-indexing approach of removing the modulus from the sub-index (Q_i) and introducing a positive index (PI) and a negative index (NI) for each water sample. The cumulative sum of the PI and NI was computed for grading the water quality at a sampling location as m-HPI. Also, the upper limit of the PI (U_L) was established for assessing the suitability of a water sample for drinking by assigning a value of the upper permissible concentration to the monitored parameter ($M_i = S_i$). The sub-indexing approach for the method was as follows:

$$Q_i = \sum_{i=1}^{n} \frac{(M_i - I_i)}{I_i} \tag{6.6}$$

It is noteworthy that this particular method suffers from the phenomenon of eclipsing where the relatively high value of one or more water quality parameters may not have a significant influence on the PI consequently estimating its final end result lower than U_L. In other words, if the concentrations of one or two heavy metals are very high and the concentrations of the remaining heavy metals are very low, the water quality would be graded very good or good ($PI < U_L$) considering it suitable for drinking purpose, when in reality it is not.

To address the shortcomings of the existing indices, Entropy-weighted Heavy Metal Contamination Index (EHCI) has been proposed. The index comprises the subsequent steps:

- Data pre-treatment: The first step involves the normalization of each monitored heavy metal in a particular sample so as to eliminate errors caused by different quantity grades of quality.

$$v_{ij} = \frac{c_{ij} - (c_{ij})_{min}}{(c_{ij})_{max} - (c_{ij})_{min}} \tag{6.8}$$

  Where v_{ij} is construction function of normalization for an evaluated parameter (n) in a particular water sample (m).

- Determination of entropy weights: Information entropy was adopted in determining the weight of each heavy metal (w_j) as per the following equation (Pei-Yue et al., 2010):

$$w_j = (1 - E_j) / \sum_{j=1}^{t} (1 - E_j) \tag{6.9}$$

$$E_n = -(1/\ln m) \sum_{i=1}^{m} T_{ij} \ln T_{ij} \tag{6.10}$$

  Where, T_{ij} is the probability of occurrence of j^{th} parameter in i^{th} water sample and was computed by:

$$T_{ij} = \frac{v_{ij}}{\sum v_{ij}} \tag{6.11}$$

- Sub-indexing: The description of the water quality called for sub-indexing of the water quality parameters. Sub-indices can be classified as absolute and relative indices. Absolute indices are independent on water quality standards while relative indices are dependent on water quality standards (Swamee and Tyagi, 2000). Relative sub-indexing approach has been employed in the present study. Sub-index for each parameter was assigned by:

$$Q_j = \left(\frac{C_j}{S_j}\right) * 100 \tag{6.12}$$

- Aggregation of sub-indices and development of EHCI: Weighted arithmetic mean method has usually been employed by earlier researchers. The weighted arithmetic mean method has been employed in the present study for the computing the EHCI as per the following equation:

$$EHCI = \sum_{j=1}^{n} w_j Q_j \tag{6.13}$$

Where Cj is measured concentration of the parameter

The quality grading scale of the proposed index suggested by Jian-Hua et al. (2011) designates waters with EHCI < 50 to be of "excellent" quality, EHCI between 50-100 as "good", EHCI between 100-150 as "average", EHCI between 150-200 as "poor" and EHCI > 200 as "extremely poor". EHCI was applied on the observed heavy metal data sets. Variation of EHCI along the stretch of rivers has been depicted in Fig. 6.2 (a – h).

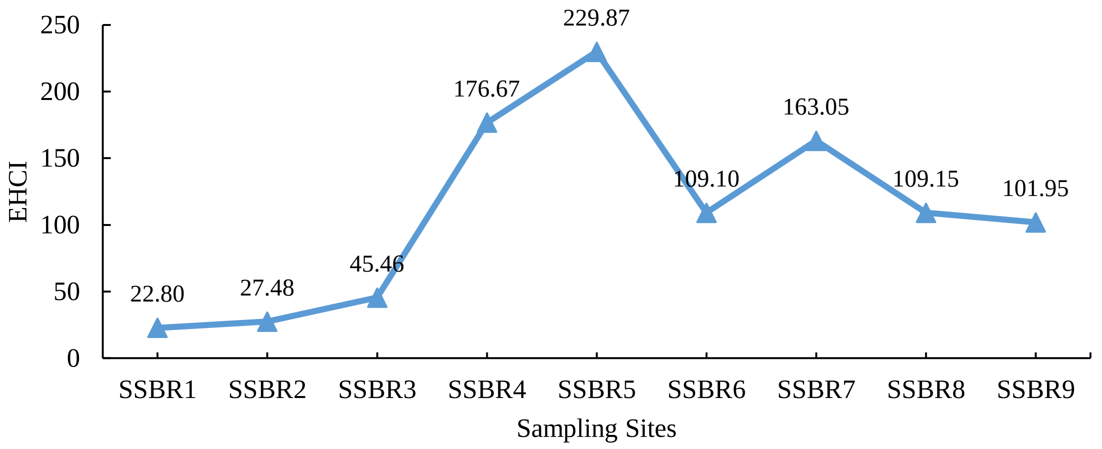


Fig. 6.2 (a). Spatial variation of EHCI in Baralia River

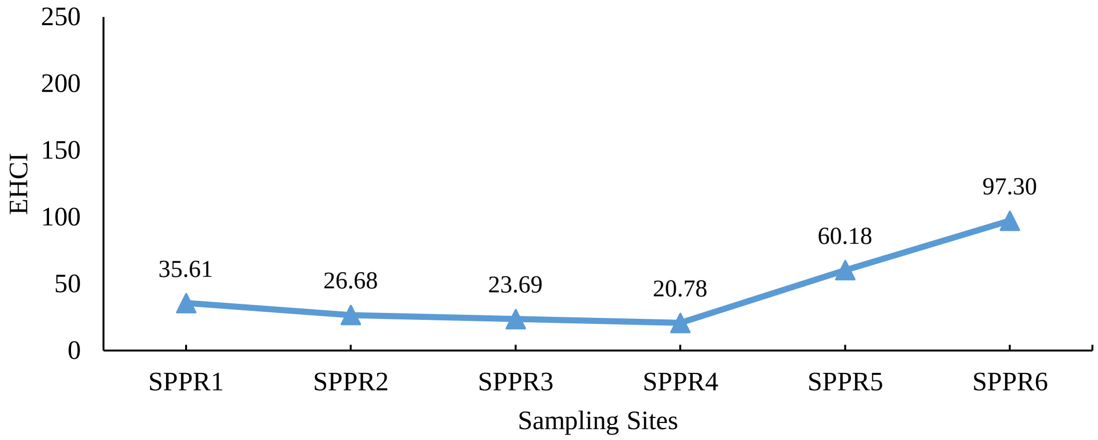


Fig. 6.2 (b). Spatial variation of EHCI in Puthimari River

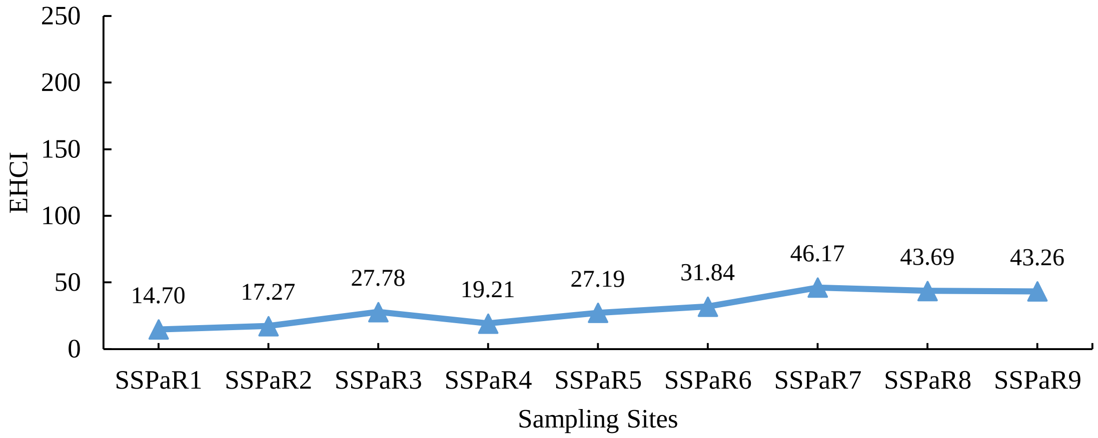


Fig. 6.2 (c). Spatial variation of EHCI in Pagladia River

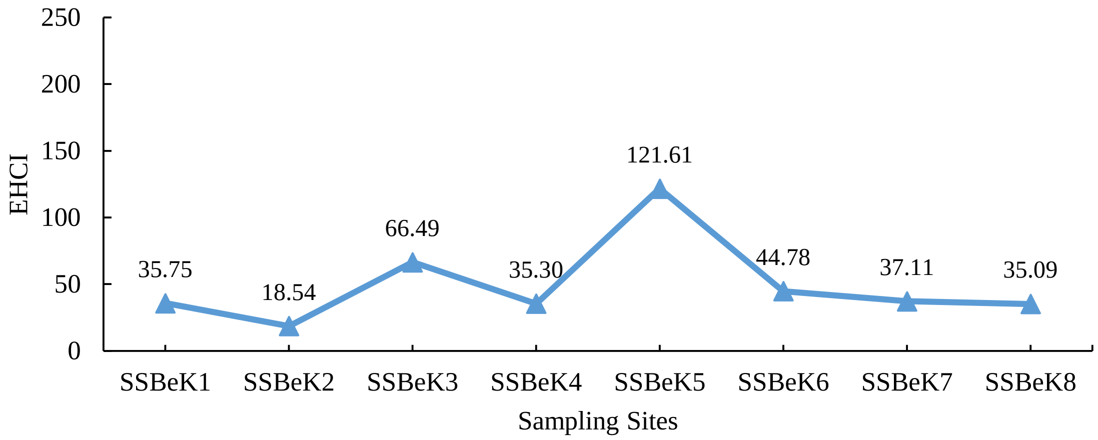


Fig. 6.2 (d). Spatial variation of EHCI in Beki River

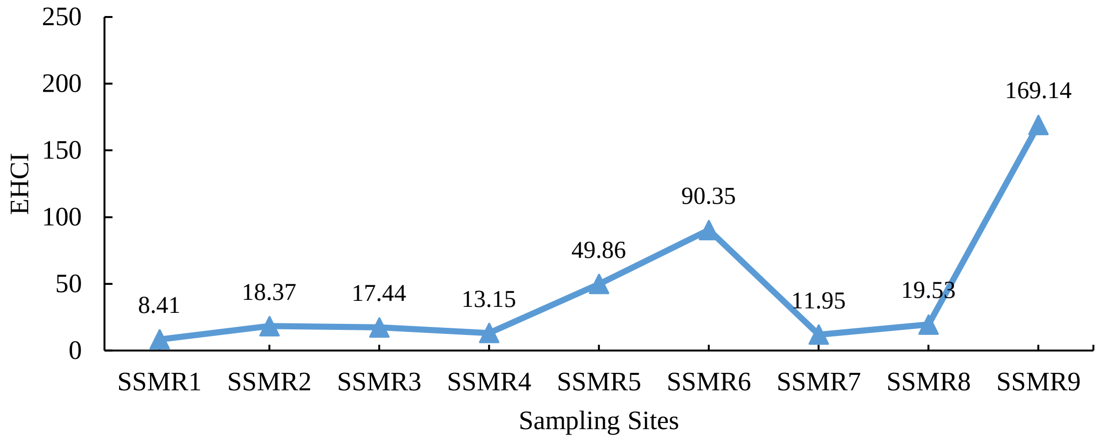


Fig. 6.2 (e). Spatial variation of EHCI in Manas River

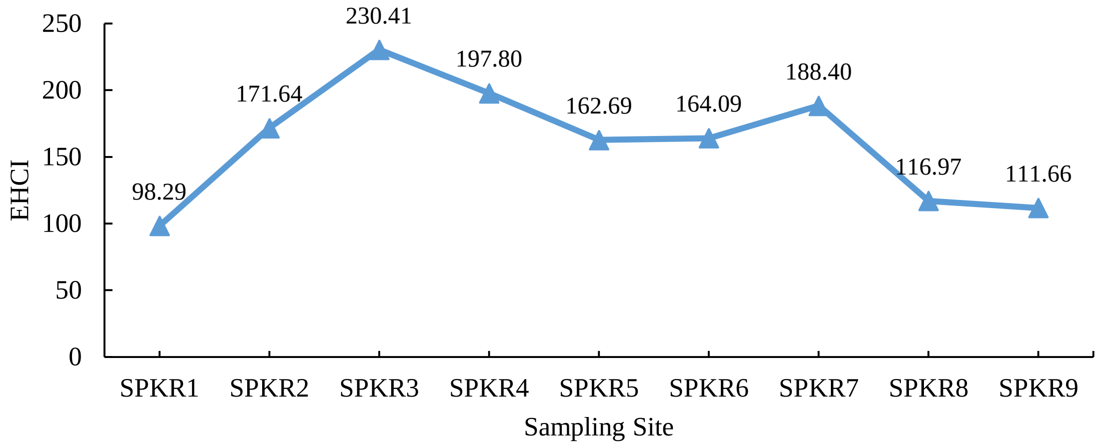


Fig. 6.2 (f). Spatial variation of EHCI in Kolong River

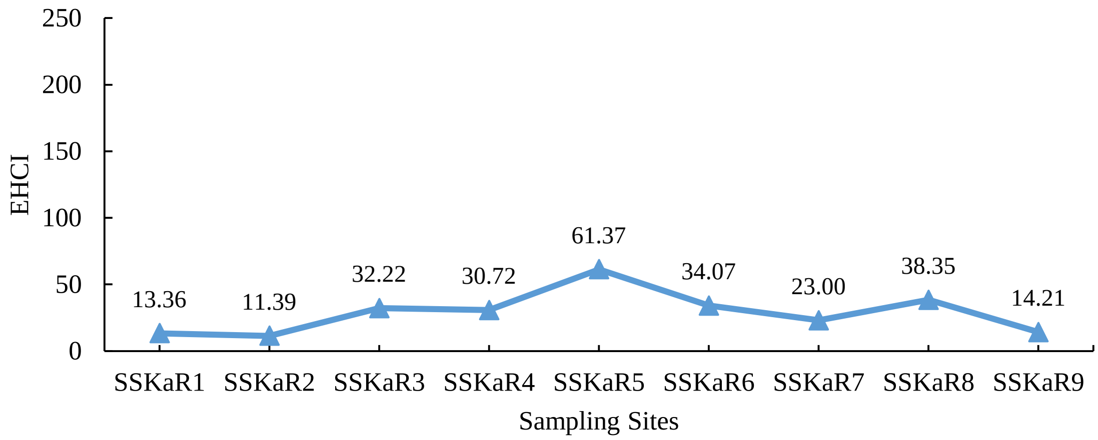


Fig. 6.2 (g). Spatial variation of EHCI in Kameng River

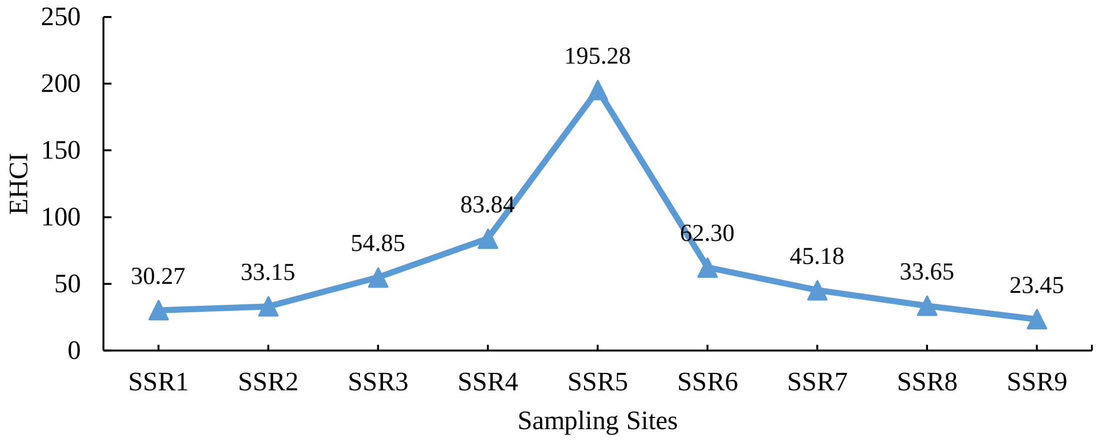


Fig. 6.2 (h). Spatial variation of EHCI in Brahmaputra River

6.3 DEVELOPMENT OF ENTROPY WEIGHTED IRRIGATION WATER QUALITY INDEX (EIWQI)

For development of WQIs, it is essential to select the few important water quality parameters which are normally accepted as informative for a desired end use and represent the overall water quality (Misaghi et al., 2017). To the best of the authors' knowledge, Misaghi et al., (2017) introduced first water quality index for the assessment of river water for irrigation purposes based on National Sanitation Foundation Water Quality Index (NSFWQI). In this index, water quality parameters were amended to account for parameters suitable for irrigation and weighting chart was generated according to the FAO 29 guideline. Misaghi et al., (2017) considered seven parameters (SAR, Na^+, pH, Cl^-, HCO_3^-, EC and TDS). Weight of the selected parameters was determined by opinion of experts and analytical hierarchy process (AHP). This index suffers from two major drawbacks. Based on literature review, it was observed that water quality for irrigation purpose generally evaluated using some derived parameters such as permeability index (PI), Kelly's ratio (KR), magnesium adsorption ratio (MgR), sodium adsorption ratio (SAR), soluble sodium percentage (SSP), and residual sodium carbonate (RSC) (Chung et al., 2015; Gautam et al., 2015; Thakur et al., 2016; Falah and Haghizadeh, 2017; Rahman et al., 2017; Salifu et al., 2017; Sharma et al., 2017; Thapa et al., 2017). Misaghi et al. neglected these parameters in their index. Second drawback is determination of weight using AHP and expert's opinion. As mentioned earlier, these methods are time consuming and highly subjective. Utilizing the benefits of entropy weight, entropy weighted irrigation water quality index (EIWQI) has been proposed. EIWQI is similar to EWQI and EHCI, only parameters were amended. In this index, TDS, KR, MgR, SSP, SAR, PI and RSC were used for development of EIWQI. Mathematical expressions for calculation of KR, MgR, SSP, SAR, PI and RSC are given in ‘Materials and Methods’ chapter (Eq. 3.12-3.17). KR determine the hazardous effect of sodium on water quality for irrigation usage. Although, Mg^{2+} is an essential constituent for healthy growth of plant, excessive amount of Mg^{2+} in irrigation water can be lethal to plants as it leads to reduced availability of K^{+} in soil where Mg^{+2} concentrations are raised. If MgR <50, then the water is safe and suitable for irrigation (Khodapanah et Al, 2009). SSP represents sodium hazard. SSP more than 50% may lead to Na^{+} buildups that cause breakdown in the soil’s physical properties (Ayers and Westcot 1985; Alobaidy et al. 2010). SAR is sodium adsorption by soil, and it evaluates the relative proportion of Na^{+} in water with respect to Ca^{+2} and Mg^{+2}. A high level of Na^{+} in water increases soil alkalinity but reduces soil permeability along with availability of water to plants (Nagarajah et al. 1988; Arveti et al. 2011). PI evaluates the effects of long-term use of irrigation water that contains high amount of Na^{+}, Ca^{+2}, Mg^{+2}, and HCO_3^{-} on permeability of soil (Davraz and O¨ zdemir 2014; Li et al. 2016a; Vasanthavigar et al. 2012). RSC indicates the hazardous effect of carbonate and bicarbonate to soils and plants.

Spatial variation of water quality in terms of EIWQI has been shown in Fig. 6.3 (a – h). This index can be helpful for identifying ideal locations along the river stretch with water quality pertaining to irrigation standards and can be considered as a major criterion for constructing intake structures for irrigation canals.

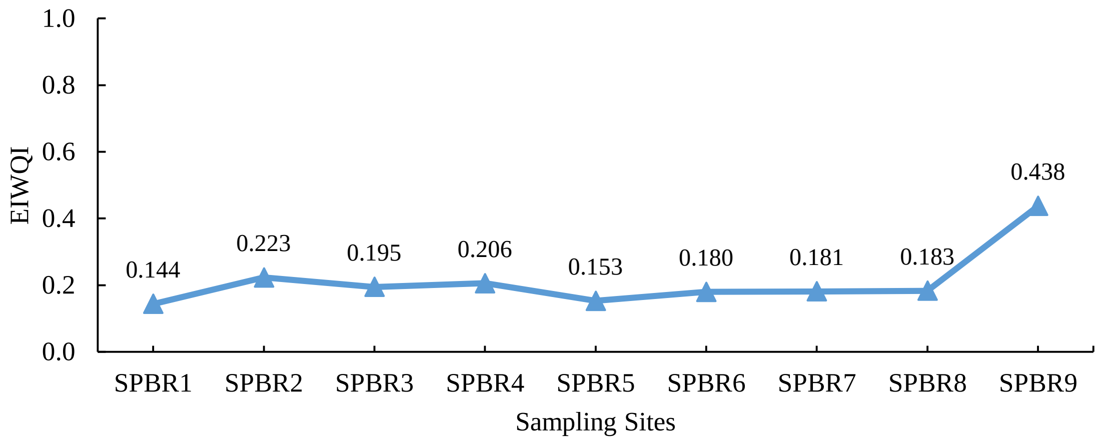


Fig. 6.3 (a). Spatial variation of EIWQI in Baralia River

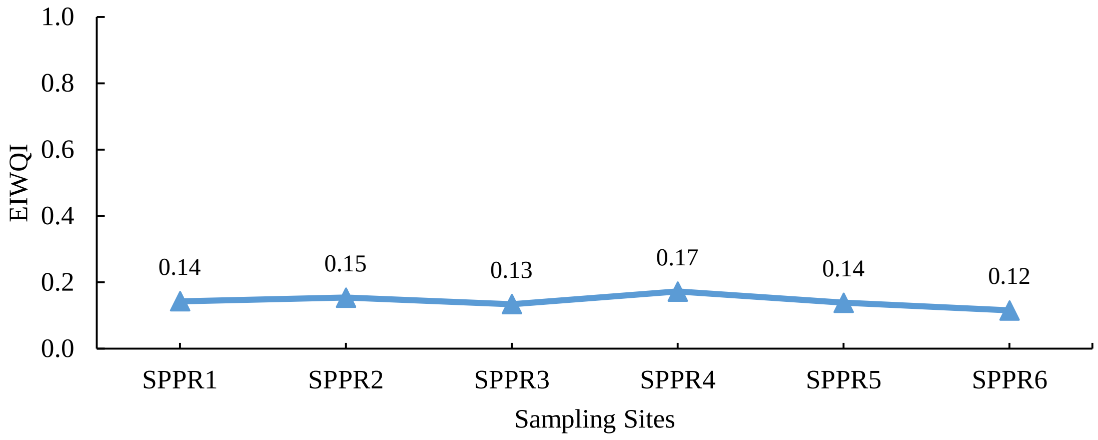


Fig. 6.3 (b). Spatial variation of EIWQI in Puthimari River

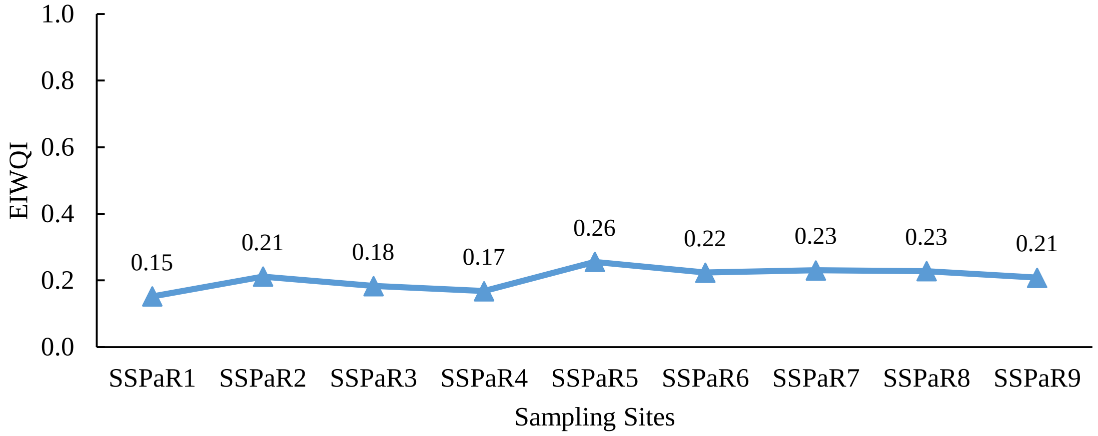


Fig. 6.3 (c). Spatial variation of EIWQI in Pagladia River

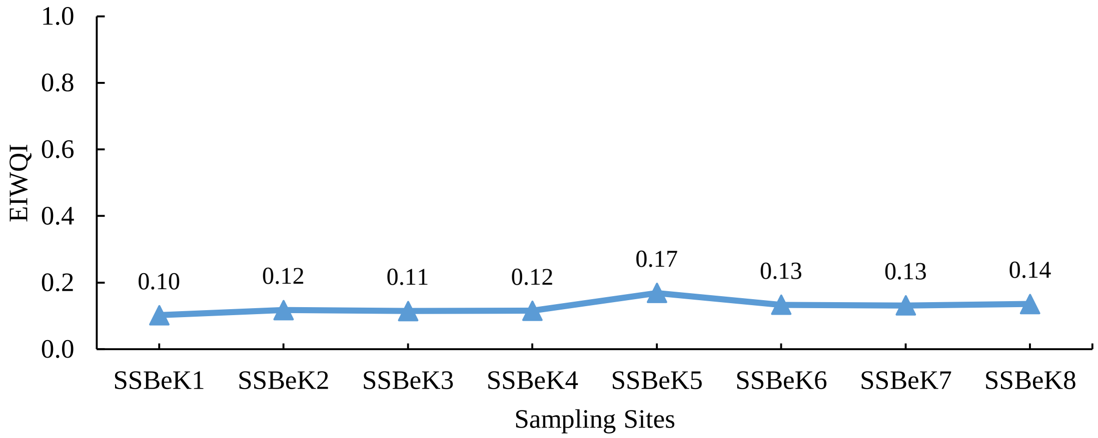


Fig. 6.3 (d). Spatial variation of EIWQI in Beki River

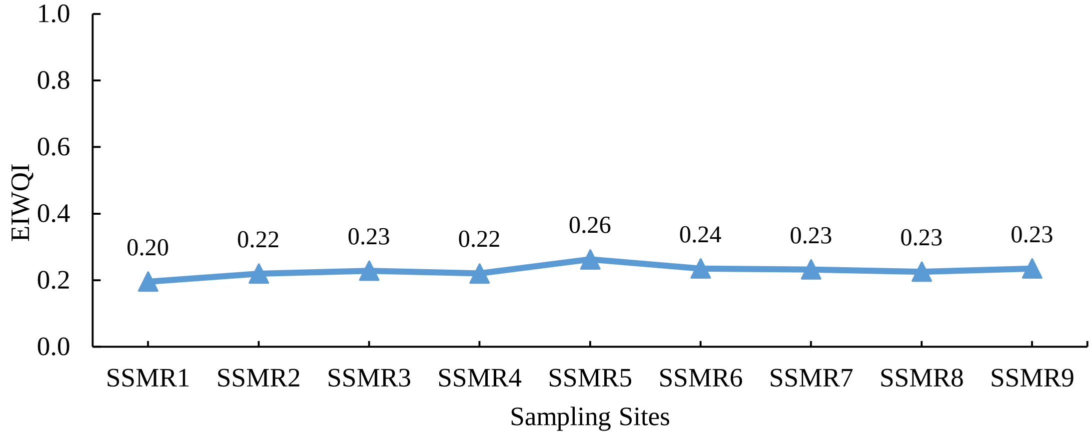


Fig. 6.3 (e). Spatial variation of EIWQI in Manas River

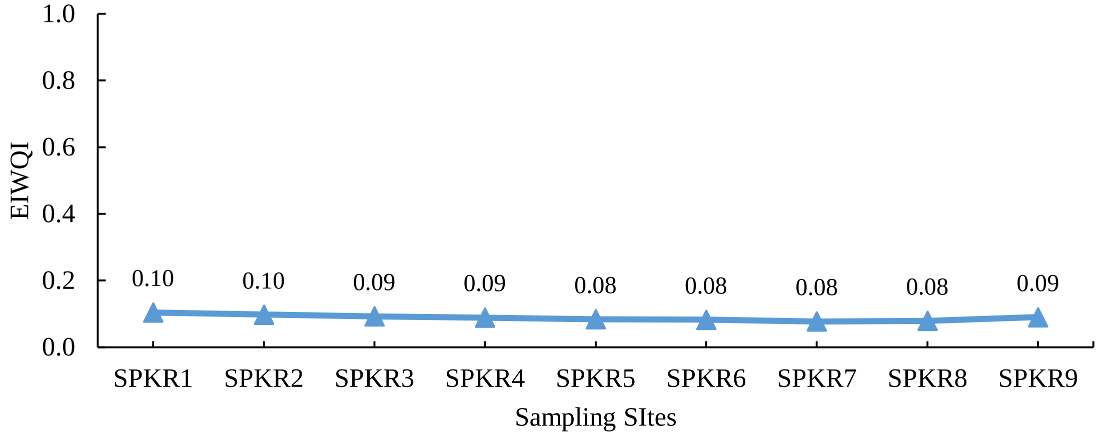


Fig. 6.3 (f). Spatial variation of EIWQI in Kolong River

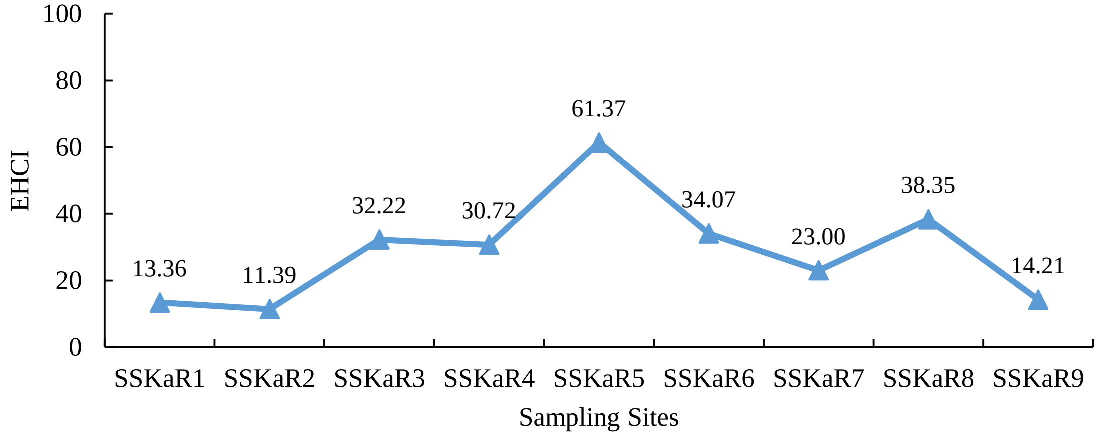


Fig. 6.3 (g). Spatial variation of EIWQI in Kameng River

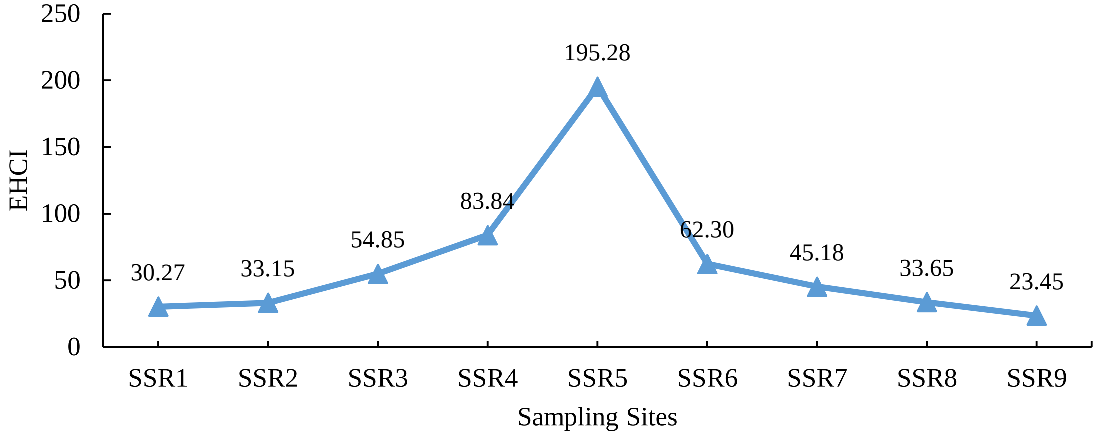


Fig. 6.3 (h). Spatial variation of EIWQI in Brahmaputra River

6.4 OVERALL RANKING OF SAMPLING SITES

EWQI evaluated the surface water quality with respect to physicochemical parameters while EHCI evaluated the water quality with respect to heavy metals. In order to develop an insightful knowledge about the overall pollution levels, the TOPSIS methodology was applied on all the observed water quality parameters (physicochemical and heavy metals). The TOPSIS methodology which is a reliable method utilizing rough set theory results in the analysis from different parameter weights. TOPSIS was used to develop overall ranking of sampling sites, such that the lowest TOPSIS rank in river would indicate the most polluted sampling sites. The ranks are important to serve as entities on which policy making and restoration of the rivers may be prioritized. The TOPSIS ranks of the sampling sites of the rivers have been shown in Fig. 6.4 (a-h).

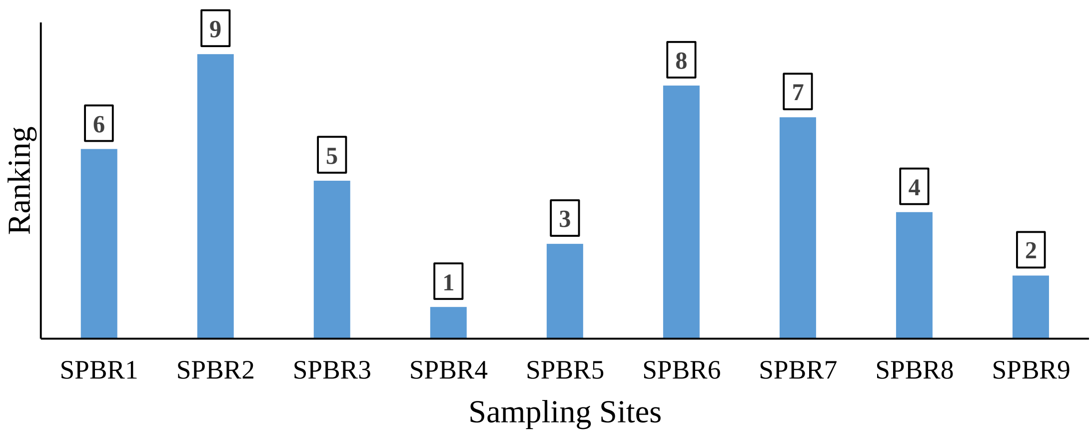


Fig. 6.4 (a). TOPSIS rank of all sampling sites of Baralia River

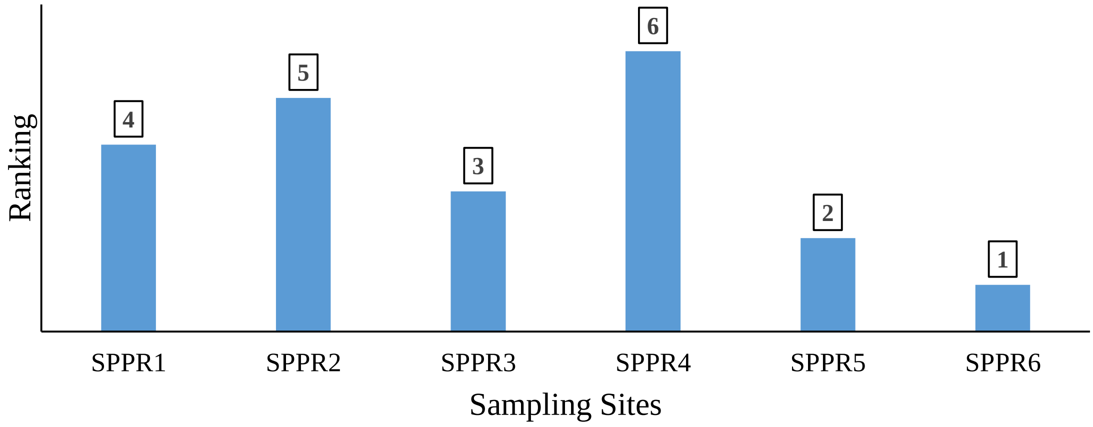


Fig. 6.4 (b). TOPSIS rank of all sampling sites of Puthimari River

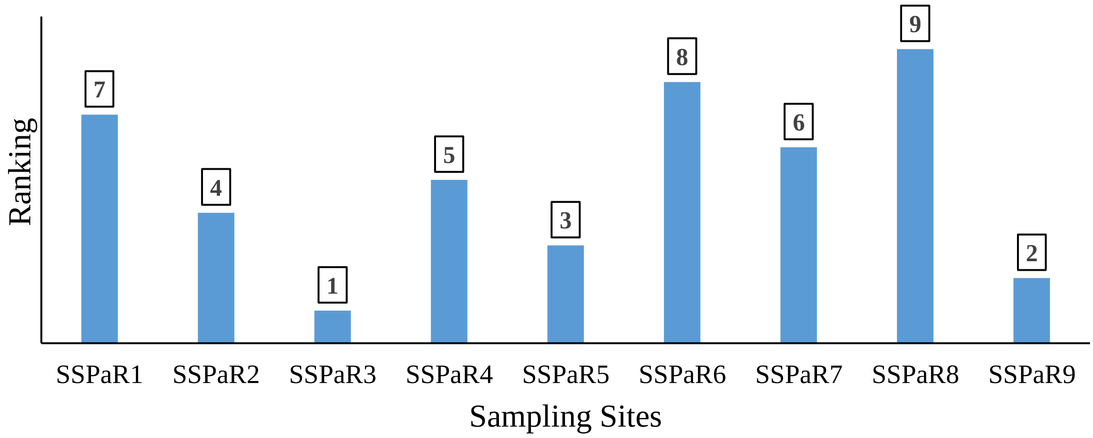


Fig. 6.4 (c). TOPSIS rank of all sampling sites of Pagladia River

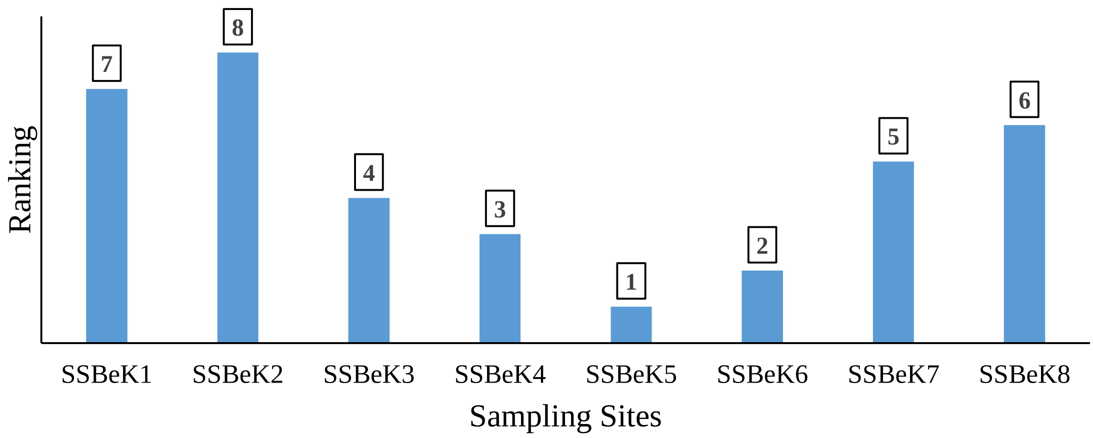


Fig. 6.4 (d). TOPSIS rank of all sampling sites of Beki River

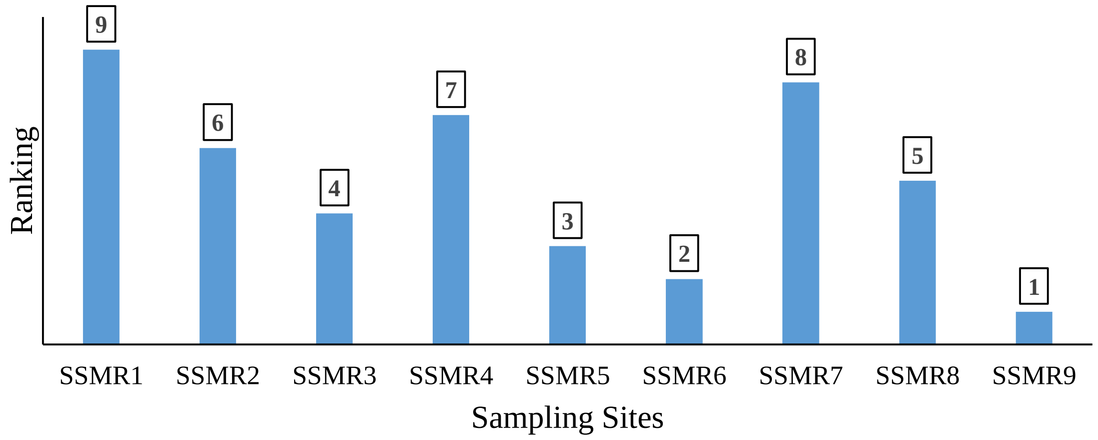


Fig. 6.4 (e). TOPSIS rank of all sampling sites of Manas River

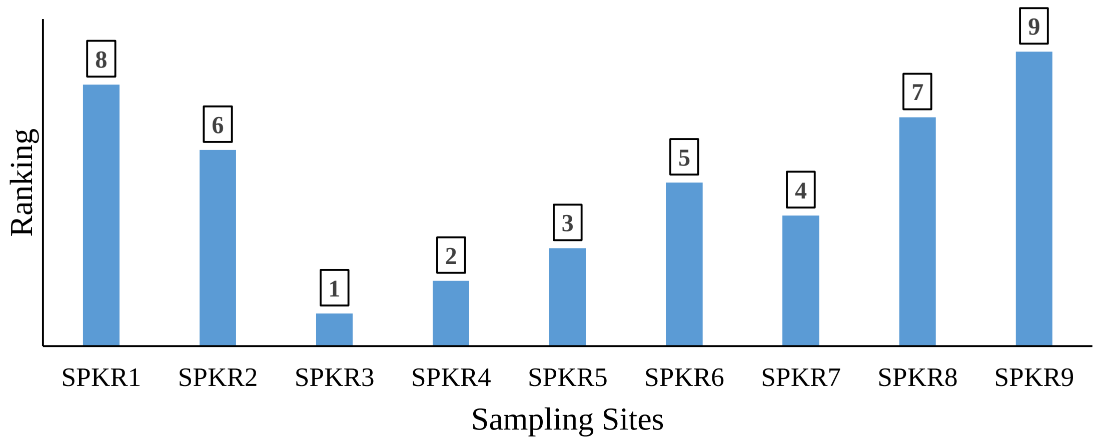


Fig. 6.4 (f). TOPSIS rank of all sampling sites of Kolong River

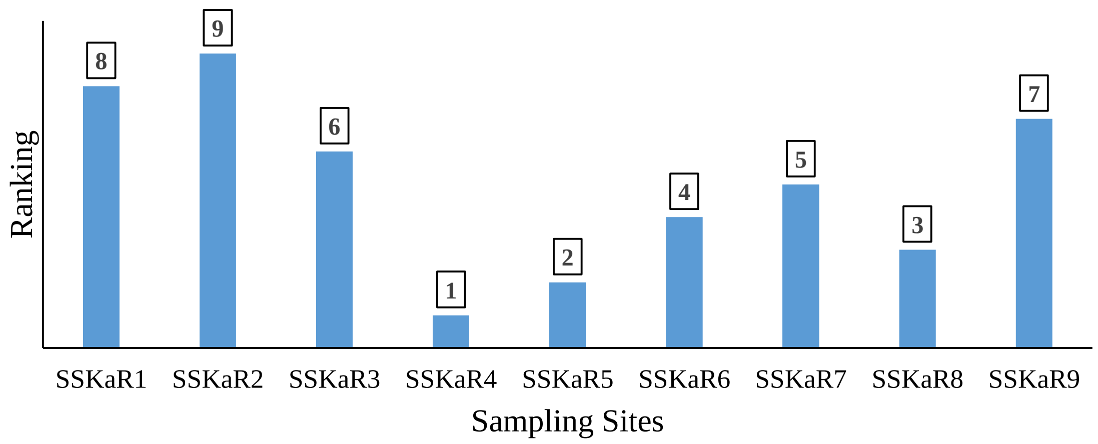


Fig. 6.4 (g). TOPSIS rank of all sampling sites of Kameng River

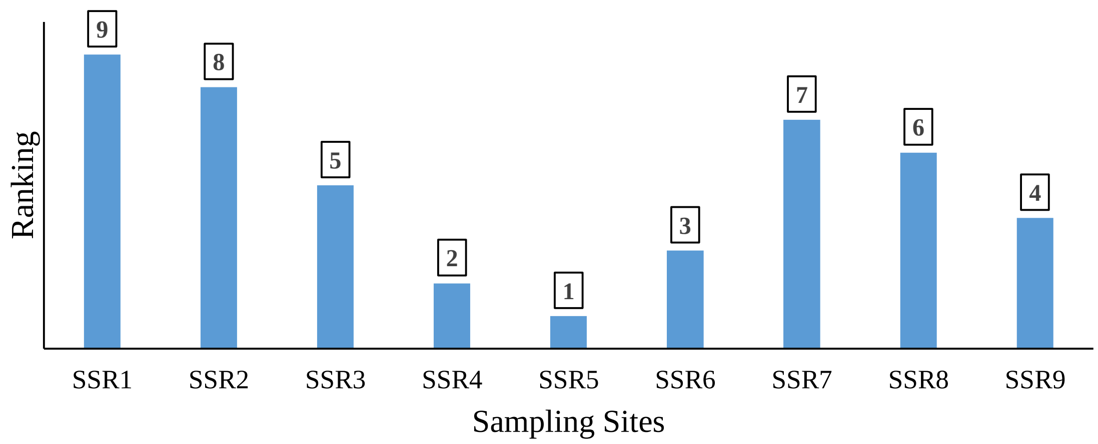


Fig. 6.4 (h). TOPSIS rank of all sampling sites of Brahmaputra River

6.5 CHAPTER CONCLUSION

In this 3rd phase of the study, an attempt was carried out to characterize the surface water quality for various purposes by using the information entropy method. The important conclusions from the study were drawn as follows:

- EWQI categorized the water quality at all sampling sites (except SPBR3, SPBR4, SSBeK5 and SSR5) as "excellent" or "good" with respect to drinking water quality standards. Water quality at sampling sites SPBR3, SPBR4, SSBeK5 and SSR5 was found to be in "poor" category.
- EWQI classification provides the more reasonable results than conventional WQIs due to its simplicity, accuracy, and ignoring of artificial weight.
- For heavy metals, proposed index EHCI overcome the shortcomings of the existing indices - HPI, CI, HEI and mHPI.
- Similarly, an index EIWQI was suggested for the evaluation of water quality variability for irrigation.
- Finally, an overall ranking has been assigned to all sampling sites using TOPSIS which incorporates all the observed datasets.

7

Chapter

Identification of ideal monitoring locations

This chapter covers

- Spatial variation of EWQI in Deepor Beel
- Spatial variability of EHCI in Deepor Beel
- Geospatial analysis of Water Quality variability Deepor Beel
- Risk assessment of heavy metals on human health
- Conclusions

7.1 IDENTIFICATION OF IDEAL MONITORING LOCATIONS

In recent decades, anthropogenic influences leading to water quality degradation, a consequence of industrialization and urbanization has created the need of efficient monitoring programs for undertaking better policies and stream restoration efforts (Strobl and Robillard, 2008; Behmel et al., 2016). These monitoring programs primarily focus on the aggregation and interpretation of water quality variables requiring novel techniques for identifying the cumulative effect of point sources and non-point sources. Also, the variability of precipitation has a considerable impact on the dilution of pollution loads in aquatic bodies (Zhao et al., 2015). Generally, variability may be defined as the state of unevenness or lack of uniformity over different scales. With increased effects of point sources, non-point sources and precipitation on the water quality, it becomes a challenging task to investigate the spatial and temporal variability of the water quality of a waterbody which might also be an indicator of the addition of new sources or retreat of existing ones.

Spatial variability is characterized by different values of observed attributes at different geographical locations while temporal variability is focused on computing the disorder of an attribute over a temporal scale (Mishra et al., 2009). Although variance is the most common statistic used to assess variability, several studies have reported that information entropy might serve as a better measure as it measures the dispersion of the probability density function of an attribute unlike variance which measures concentration metrics of an attribute around the mean (Mishra et al., 2009; Silva et al., 2016). The application of variance based dispersion indices have been observed in the frequency distribution of coliforms and benthic macroinvertebrates (Christian and Pipes, 1983; Moulton et al., 2000). As a matter of fact, interpretation of variance as a measure of disorder should be done by accounting for its restraints (Soofi, 1997).

Most of the studies employing information entropy in the assessment of space-time variability has been limited to hydro meteorological variables such as rainfall and streamflow in arid and semi-arid regions, consequently accounting the cumulative sources of such variability (Amorocho and Espildora, 1973; Sivakumar, 2001; Hsu et al., 2006; Mishra et al., 2009; Brunsell, 2010; Liu et al., 2013; Silva et al., 2016; Roushangar and Alizadeh, 2018;). Assessment of space-time variability of rainfall for water resources management has been assessed in the light of information entropy (Mishra et al., 2009; Zhang et al., 2016). Subsequently, such studies were extended to streamflow and watershed topography effects in arid and semi-arid regions (Djebou et al., 2014; Silva et al., 2016). On the basis of such assessments, it is crucial to align water quality monitoring and assessment studies for better policy and decision making.

In investigation of entropy-based variability in water quality, a plausible approach may be starting with the aggregation of large and complex datasets of water quality variables generated over a period of time by monitoring programs into numerical scores known as Water Quality Indices (WQIs). A disorder index (DI) computed on the principle of maximum entropy (POME) and therefore, quantifying the variability among these WQIs over a time series scale at each sampling location would be capable of representing the cumulative variability of water quality at that particular location. DIs have significant benefits in (i) quantifying variability of a data set over a temporal scale (ii) reporting a minimum value when all the values in the data sets are same (iii) depicting less series variability for narrow ranges of values (Martino et al., 2012). Furthermore, the WQI is just a numerical score and is bound to change even with a minor change in water quality variables or sub-indexing, and is also prone to phenomena such as eclipsing and exaggeration (Swamee and Tyagi, 2000). The DI quantifying the variability among WQIs at a particular location over a monitoring period by a numerical score would be effective in identifying sites in an area which have undergone high and low variability of water quality due to cumulative effects of discharge of sewage and surface run-off, dilution of wastewaters and change in water depth with the aid of geospatial techniques and isoinformation lines (Mishra et al., 2009; Djebou et al., 2014).

The present study attempts to investigate upon the variability of water quality in the Deepor Beel (Fig.7.1). Being a natural wetland, it has been host to biologically diverse ecosystems with species richness in floral and faunal diversity. However, in recent years due to massive encroachment, the Beel has shrunk from an original area of 4000 ha to 700 ha (Kapil and Bhattacharyya, 2012). Despite being listed in the Directory of Asian Wetlands, and its international importance as a Ramsar site, the dumping ground of Guwahati city was assigned adjacent to the Beel in the year 2008 (Choudhury and Gupta, 2017). The surrounding low lying area of the dumpsite connects it with the Deepor Beel. Although, several studies have focused on the solid waste of the dumpsite and its impact on the water quality of the Beel (Kapil and Bhattacharyya, 2012; Choudhury and Gupta, 2017), no study exists upon the variability of the water quality in the total area of the Beel due to combination of all possible natural and anthropogenic influences. Consequently, this study aims to provide a transparent picture in quantifying the variability of water quality in Deepor Beel with the aid of entropy based diversity indices.

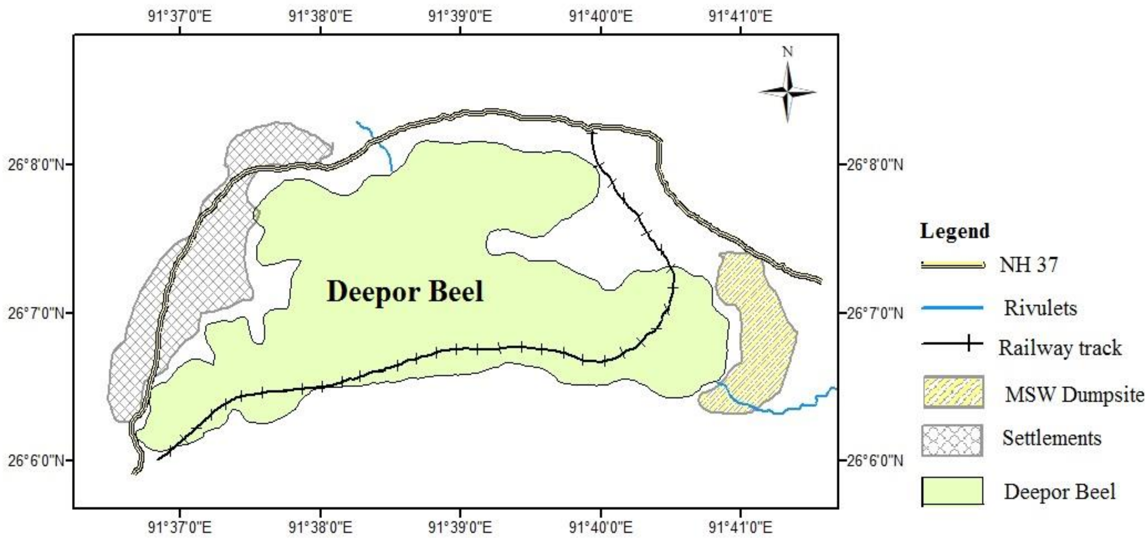


Fig. 7.1. Deepor Beel

The study area has been essentially divided into four major sites as shown in Fig. 7.2. The site 1 identifies locations situated near the MSW dumping ground, kilns, and a host of anthropogenic influences from roadside activities. The site 1 also receives effluents from rivulets of Bharalu and Basistha-Bahini rivers. These rivers also carry sewage effluents of Guwahati city and discharge them into the wetland. The site 2 identifies locations situated near small industries and settlements in the periphery of the Beel. The site 3 identifies locations situated near the rivulet of Khonajan which connects the Brahmaputra river to the Beel and also is influenced by the roadside activities of National Highway 37 (NH 37) and small industries and manufacturing units located on it. The site 4 is essentially the midpoint of the wetland affected least by the anthropogenic influences of the wetland.

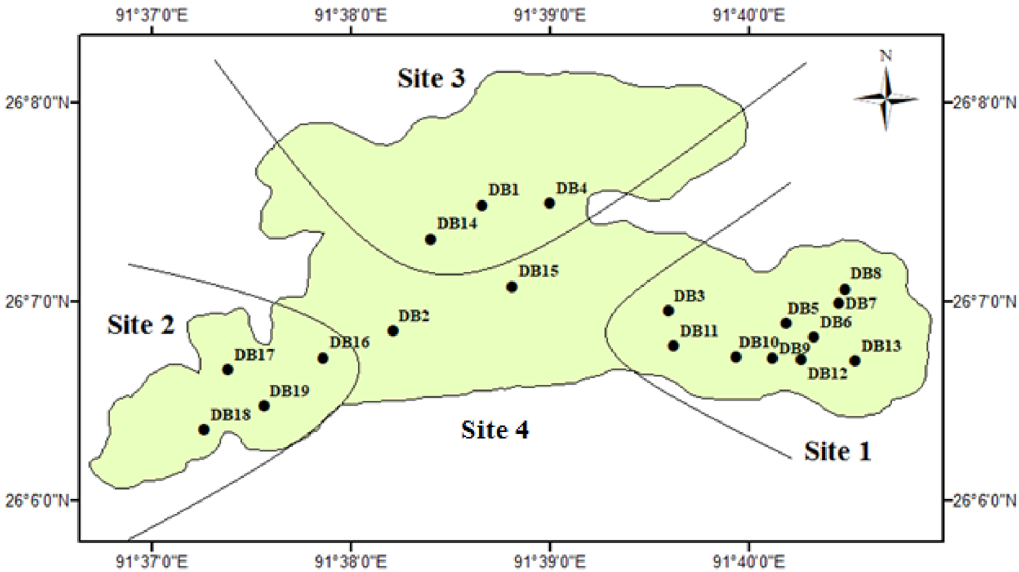


Fig. 7.2. Location of sampling sites in Deepor Beel

For this phase of the study, water samples were collected thrice in each season of all four seasons namely rainy (May–August), Post-monsoon (September–November), winter (December–February) and summer (March–April) from nineteen locations of Deepor Beel in replicates throughout the year 2017–2018. All the samples were analysed for various physicochemical parameters and heavy metals. Site wise maximum (Max), minimum (Min) and average value (Avg) of physicochemical parameters with standard deviations (SD) have been depicted in the Table 7.1

Table 7.1. Statistical summary of physico-chemical parameters

Sampling Site		pH	EC	Turbidity	TH	TA	DO	BOD	COD	TDS	Na^{+}	K^{+}	Mg^{2+}	Ca^{2+}	F^{-}	Cl^{-}	$(SO_4)^{2-}$	$(NO_3)^{-}$
DB1	Max	7.45	0.26	40.00	72.00	74.00	12.50	36.00	122.88	143.33	6.70	10.83	4.93	94.57	0.41	22.44	13.64	0.44
	Min	6.92	0.19	20.40	54.00	56.00	9.81	8.70	55.17	53.97	0.58	0.09	0.29	24.17	0.00	13.95	7.01	0.01
	Avg	7.22	0.22	28.00	59.50	64.50	11.02	25.05	100.27	115.99	3.84	4.18	2.22	48.36	0.26	18.36	10.90	0.22
	SD	0.22	0.03	9.41	8.54	7.55	1.24	13.39	30.80	41.71	3.22	5.07	1.99	32.16	0.19	3.47	2.80	0.19
DB2	Max	7.87	0.24	70.00	64.00	68.00	11.50	32.40	368.64	303.33	7.10	10.34	3.91	83.54	0.74	21.48	60.91	2.25
	Min	6.50	0.20	17.00	48.00	56.00	8.53	8.40	87.27	56.10	0.21	0.07	0.32	36.40	0.30	12.82	10.76	0.06
	Avg	7.19	0.21	37.75	56.50	61.50	10.34	19.63	208.48	222.36	2.86	4.00	1.67	54.85	0.44	16.62	27.64	0.72
	SD	0.57	0.02	22.68	6.61	5.00	1.38	10.61	124.24	112.97	3.12	4.91	1.63	22.09	0.21	3.86	22.70	1.02
DB3	Max	7.53	0.29	23.40	60.00	80.00	10.24	35.40	174.55	790.00	5.80	10.68	0.99	93.84	0.36	18.92	23.43	2.35
	Min	6.60	0.22	14.80	54.00	58.00	4.37	9.30	27.60	39.20	0.10	0.20	0.35	21.97	0.13	11.36	7.12	0.15
	Avg	7.05	0.25	20.38	56.00	68.50	8.02	20.93	86.91	281.55	2.77	4.17	0.66	49.42	0.26	15.16	12.17	0.79
	SD	0.44	0.03	3.86	2.83	9.15	2.81	10.78	64.17	342.86	2.73	4.94	0.36	32.14	0.11	3.24	7.62	1.04
DB4	Max	7.43	0.27	17.00	84.00	88.00	11.20	29.40	116.36	866.67	4.60	8.07	6.85	93.88	0.44	19.19	25.64	0.65
	Min	6.54	0.20	12.80	48.00	62.00	8.00	11.10	26.67	40.97	0.61	0.07	0.23	27.87	0.18	14.02	4.45	0.00
	Avg	6.97	0.23	14.78	62.50	70.50	8.96	20.33	71.81	304.49	2.35	3.43	4.53	62.65	0.31	17.19	14.96	0.36
	SD	0.40	0.03	1.80	15.78	12.26	1.50	10.14	37.64	378.92	1.93	3.97	3.04	31.74	0.11	2.36	8.86	0.33

DB5	Max	6.85	0.26	16.50	98.00	74.00	8.39	33.60	153.60	166.67	4.16	9.74	1.46	112.05	0.42	21.55	15.18	4.47
	Min	6.67	0.23	6.10	46.00	56.00	4.17	8.70	13.33	44.30	0.04	0.17	0.28	24.88	0.15	11.85	5.69	0.28
	Avg	6.74	0.25	11.55	62.50	68.00	5.25	22.43	55.91	117.66	2.09	4.02	0.65	52.03	0.31	16.38	11.61	1.58
	SD	0.08	0.01	4.41	24.24	8.16	2.09	12.49	65.52	51.96	2.24	4.48	0.55	41.06	0.12	4.09	4.42	1.95
DB6	Max	6.86	0.26	26.20	58.00	94.00	9.40	37.00	153.60	166.67	3.50	8.59	3.48	112.20	0.49	121.96	54.75	1.88
	Min	6.54	0.19	6.80	48.00	46.00	2.30	15.90	29.09	40.60	0.41	0.01	0.38	28.51	0.22	10.60	8.77	0.00
	Avg	6.73	0.22	14.25	53.00	68.50	4.41	28.83	86.36	129.23	2.07	3.65	2.15	69.13	0.34	41.94	22.07	0.51
	SD	0.14	0.03	8.79	4.16	19.82	3.34	9.85	51.17	59.73	1.61	3.96	1.29	43.42	0.13	53.45	21.86	0.91
DB7	Max	6.92	0.29	14.30	64.00	88.00	9.80	31.80	116.36	133.33	11.00	6.88	3.86	94.59	0.49	22.78	30.02	7.75
	Min	6.74	0.21	3.10	52.00	62.00	3.68	12.00	26.67	57.17	2.30	0.00	0.32	34.52	0.12	14.00	7.64	0.33
	Avg	6.84	0.26	9.18	59.00	72.00	7.81	19.58	72.59	108.46	6.68	2.90	2.51	55.17	0.31	18.19	21.81	3.15
	SD	0.08	0.04	4.66	5.03	11.89	2.82	9.43	39.63	35.30	3.76	3.30	1.59	27.62	0.17	4.75	10.60	3.49
DB8	Max	8.19	0.57	38.30	64.00	92.00	14.25	32.40	232.73	246.67	8.70	10.72	1.24	89.18	1.05	25.32	66.67	0.84
	Min	6.78	0.19	7.20	56.00	60.00	2.65	14.10	27.60	52.23	0.30	0.07	0.33	40.80	0.03	19.28	10.68	0.11
	Avg	7.22	0.32	18.18	60.00	75.00	10.96	24.53	122.47	146.39	4.15	4.90	0.59	65.20	0.47	20.92	25.87	0.52
	SD	0.65	0.17	14.06	4.62	13.22	5.55	7.76	84.47	84.91	4.05	5.34	0.44	24.94	0.42	2.94	27.25	0.35
DB9	Max	7.28	0.25	29.60	60.00	86.00	11.34	38.40	122.88	167.00	11.87	5.58	3.43	103.53	0.49	18.54	40.51	2.27
	Min	6.43	0.20	8.60	42.00	38.00	5.80	12.00	26.67	39.87	0.09	0.07	0.37	29.00	0.16	12.90	9.10	0.00
	Avg	6.75	0.23	18.50	49.00	60.00	7.38	24.30	87.17	116.72	3.86	2.89	2.45	61.04	0.29	15.14	18.83	0.97
	SD	0.37	0.02	9.78	7.75	20.00	2.66	12.56	44.00	55.21	5.40	3.10	1.41	36.13	0.15	2.46	14.85	1.00

DB10	Max	6.94	0.26	25.00	62.00	78.00	7.12	40.20	240.00	153.33	6.85	9.15	2.68	86.84	0.87	22.32	22.38	2.29
	Min	6.67	0.21	10.60	50.00	68.00	2.76	10.80	55.17	40.50	0.74	0.14	0.30	24.00	0.22	11.82	7.92	0.00
	Avg	6.80	0.23	18.15	54.00	73.00	5.77	23.40	156.64	91.87	3.03	3.73	1.86	52.32	0.45	16.10	14.42	1.07
	SD	0.14	0.02	5.91	5.42	4.76	2.02	14.41	86.17	52.43	2.65	4.34	1.11	29.39	0.29	4.62	5.98	0.94
DB11	Max	7.20	0.27	22.50	60.00	90.00	10.00	30.60	174.55	157.00	11.50	10.62	3.11	82.92	0.77	17.40	26.65	1.43
	Min	6.65	0.20	12.80	42.00	52.00	3.82	10.20	26.67	36.13	0.04	0.01	0.27	10.84	0.25	15.48	8.62	0.00
	Avg	6.93	0.23	18.55	50.00	68.00	8.07	21.00	90.59	109.95	4.30	4.26	1.93	36.63	0.45	16.36	17.68	0.54
	SD	0.30	0.03	4.59	7.48	16.08	2.92	10.22	62.19	52.84	5.05	4.79	1.32	31.86	0.23	0.86	7.40	0.62
DB12	Max	6.85	0.40	19.20	56.00	80.00	7.06	44.00	145.45	190.00	6.10	10.01	2.80	91.95	0.52	19.02	72.80	2.25
	Min	6.32	0.18	4.00	42.00	52.00	2.70	9.60	26.67	55.50	0.04	0.03	0.33	30.37	0.29	12.82	9.79	0.00
	Avg	6.60	0.28	12.85	49.00	69.00	3.95	22.93	85.63	128.04	3.14	4.21	1.50	52.30	0.41	15.64	28.22	0.72
	SD	0.23	0.09	7.05	7.02	12.06	2.09	16.11	67.56	58.44	2.97	4.85	1.07	28.08	0.10	2.55	29.84	1.03
DB13	Max	6.22	0.38	84.00	82.00	108.00	9.24	84.00	110.34	253.33	10.70	9.24	1.76	105.53	0.44	77.77	78.42	0.42
	Min	5.80	0.27	6.40	62.00	24.00	1.50	24.60	30.00	45.00	0.06	0.33	0.35	26.51	0.22	28.00	9.51	0.00
	Avg	6.11	0.33	38.65	71.00	59.00	4.62	51.15	78.12	183.83	4.77	4.74	1.07	58.56	0.29	53.30	31.35	0.22
	SD	0.21	0.05	36.85	9.59	35.38	3.80	28.59	34.42	97.46	4.86	4.68	0.61	33.39	0.10	28.05	31.71	0.20
DB14	Max	7.12	0.38	24.50	56.00	82.00	11.30	31.50	110.34	190.00	7.40	11.41	7.43	83.57	0.48	20.25	17.66	2.17
	Min	6.72	0.21	12.10	44.00	58.00	5.87	11.70	29.09	55.23	0.06	0.03	0.34	11.41	0.00	13.22	7.44	0.00
	Avg	6.95	0.27	19.75	52.50	65.50	8.82	18.53	63.55	124.64	2.53	4.33	3.74	50.41	0.30	15.76	11.65	0.67
	SD	0.20	0.08	5.71	5.74	11.12	2.32	9.07	34.09	57.44	3.37	5.32	2.90	37.48	0.22	3.33	4.90	1.02

DB15	Max	7.98	0.28	35.50	78.00	80.00	10.50	28.80	153.60	243.33	5.80	8.96	4.26	93.16	0.71	64.64	28.98	2.43
	Min	6.45	0.20	13.60	54.00	58.00	5.00	8.70	29.09	35.30	0.08	0.05	0.39	35.50	0.19	13.57	13.95	0.00
	Avg	7.07	0.23	22.85	60.50	69.00	7.81	16.00	72.80	134.66	2.87	3.88	3.08	52.53	0.41	30.50	19.73	0.86
	SD	0.66	0.04	9.25	11.70	9.87	2.25	9.08	55.16	85.10	2.98	4.03	1.81	27.23	0.22	23.22	6.73	1.10
DB16	Max	7.72	0.30	32.00	62.00	82.00	10.40	47.40	82.75	406.67	5.11	10.53	0.36	110.30	0.77	20.10	32.76	2.16
	Min	6.90	0.19	14.00	52.00	66.00	7.00	14.40	30.72	50.07	0.00	0.53	0.00	22.28	0.00	0.00	0.00	0.00
	Avg	7.30	0.23	23.55	58.50	72.50	8.80	22.88	53.55	180.77	2.46	4.45	0.22	59.33	0.31	12.50	12.70	0.62
	SD	0.45	0.05	8.55	4.43	6.81	1.54	16.35	21.69	156.48	2.78	4.77	0.16	38.45	0.35	9.41	14.08	1.03
DB17	Max	7.36	0.27	27.00	60.00	76.00	16.96	31.20	145.45	166.67	6.40	7.95	4.25	98.64	0.76	18.98	20.18	0.64
	Min	6.36	0.21	13.30	50.00	64.00	2.80	14.10	26.67	39.73	0.53	0.10	0.33	35.21	0.20	11.78	4.83	0.01
	Avg	6.75	0.24	19.25	55.00	68.50	8.79	23.33	65.29	115.77	2.42	3.55	2.56	60.80	0.42	15.41	12.45	0.25
	SD	0.43	0.03	6.18	4.16	5.74	5.98	8.85	55.84	55.86	2.74	3.73	1.65	29.95	0.25	3.73	8.64	0.27
DB18	Max	8.16	0.27	47.20	60.00	84.00	10.20	43.20	153.60	166.67	4.30	10.18	3.44	87.76	0.30	23.09	13.23	1.98
	Min	6.54	0.18	20.40	46.00	46.00	8.15	11.10	26.67	57.33	0.06	0.22	0.27	17.58	0.18	11.78	3.71	0.02
	Avg	7.07	0.22	29.90	53.50	68.00	9.09	20.10	109.02	121.92	2.29	4.11	1.94	52.51	0.22	17.14	7.57	0.53
	SD	0.74	0.04	11.94	5.74	15.92	0.86	15.49	58.02	46.25	1.91	4.57	1.31	31.79	0.05	4.89	4.41	0.97
DB19	Max	7.41	0.38	33.20	54.00	64.00	9.07	31.80	174.55	146.67	8.80	10.83	2.85	89.82	0.40	57.43	110.55	4.45
	Min	6.60	0.13	18.00	42.00	52.00	8.03	6.90	27.60	36.83	0.00	0.06	0.25	27.48	0.00	11.31	7.66	0.00
	Avg	6.92	0.22	25.15	49.00	58.00	8.53	16.05	79.23	107.46	3.51	4.14	2.01	58.86	0.24	24.79	41.35	1.14
	SD	0.35	0.11	6.38	5.03	5.16	0.49	10.87	65.16	49.08	3.76	4.99	1.21	33.70	0.17	21.89	48.36	2.21

7.2 SPATIAL VARIATIONS OF DIFFERENT PHYSICO-CHEMICAL PARAMETERS AND EWQI

. The spatial variability of EWQI at all the sampling stations of Deepor Beel for each respective season has been depicted in the Fig. 7.3. The entropy weights of each physicochemical parameter for the four seasons has been depicted in the Table 7.2. In the present study, the entropy weights of physico-chemical parameters of Deepor Beel in the rainy season was of the order: Cl^- > NO_3^- > SO_4^{2-} > K^+ > BOD > Turbidity > F^- > EC > Mg^{2+} > pH > Na^+ > Ca^{2+} > TA > DO > TDS > TH. In the post-monsoon season NO_3^- and BOD had the highest entropy weights relative to other physico-chemical parameters. In winter and summer seasons TDS and Cl^- had the highest entropy weights respectively.

The EWQI of the rainy and post monsoon seasons were found to be in the range of 89.7-463.3 and 82.5-470.7 respectively. In the state of Assam, both the seasons receive high amount of rainfall with rainfall averages approaching as high as 300 mm. The highest EWQI value of both the seasons was observed at DB13 (site 1) which was at close proximity to the low lying area of Guwahati MSW dumpsite. The major parameters contributing to such high EWQI were the high BOD and turbidity values, and low DO values recorded at the site DB13. This was in accordance with the outcome of a study conducted by Choudhury and Gupta (2017) who concluded that the inflow of surface run-off from the low lying area of the dumpsite contributed to high values of BOD and low values of DO. The water quality ranged from "good" to "extremely poor" in the vast expanse of Deepor Beel with 52.63% of the sampling locations having water quality "poor" and "extremely poor" in the rainy season and 15.8% in the post-monsoon season respectively. The winter season painted a contrasting picture with EWQI values ranging from 59.1 to 154.7. The highest value of EWQI was observed at DB2 (site 2) lying in proximity to the road leading onto the Beel and the embankment of NE Frontier railway track thus being a host to a variety of anthropogenic influences. Furthermore, Guwahati receives rainfall as low as 6mm in the winter season which ceased the contribution of surface runoff from the municipal dumpsite to the Beel. In the winter season, DB13 was graded "good". Also, relatively less consumption of water for domestic activities in the winter season decreases the pollution load in the Bharalu and Basistha-Bahini rivers. In the summer season, the EWQI values ranged from 83.3 to 235.4 grading the water quality as "good" in some locations while "extremely poor" in the others. The highest EWQI was recorded at DB13 which was probably a consequence of surface run-off as state of Assam is one of the first to receive rainfall in India. 31.57% of the sampling locations were of water quality “poor” and “extremely poor” in the summer season. The earlier study conducted by Choudhury and Gupta (2017) also suggested that the growth of invasive water hyacinth in the Beel is responsible for the prevention of oxygen transfer from the air to the water surface thus minimizing photosynthesis and reducing DO levels in some parts of the wetland.

7.3 SPATIAL VARIABILITY OF HEAVY METALS AND EHCI

The statistical summary of heavy metal concentrations in Deepor Beel have been depicted in the Table 7.3. The spatial variability of EHCI at all the sampling stations of Deepor Beel have been shown in the Fig 7.4. The heavy metal concentrations portrayed great variability in its concentration metrics at all the four sites of Deepor Beel for the entirety of the monitoring period. Furthermore, the heavy metal concentrations exceeded their prescribed drinking water limits (WHO 2011; BIS 2012) resulting in high EHCI values at majority of the locations. High concentrations of Pb were observed in the rainy and post-monsoon season credited to the increased run-off and leaching across the wetland. This was in agreement with its higher concentration in MSW of the dumpsite (Choudhury and Gupta, 2017). In an equal footing, Cu and Zn concentrations were of greater magnitude in the rainy and post-monsoon seasons. As heavy metals are of persistent and non-biodegradable character, they have a tendency of accumulating in any ecosystem. Elevated concentrations of heavy metals may be a by-product of sewage effluents, leachate and run-off entering the wetland for a considerable amount of time. This fact is further supported by the existing literature which suggest that the heavy metal concentrations in the wetland have only increased every year (Karak et al., 2013; Choudhury and Gupta, 2017). In reference to the EHCI values, the Fig. 7.3 illustrates that majority of the peaks were observed in the rainy season. In the rainy season, the EHCI ranged between 42.4 and 488.6. 84.2% of the samples collected had “poor” and “extremely poor” water quality. The highest EHCI was recorded at DB13. However, the wetland was most contaminated in the winter season with all the locations of the Beel having water quality “poor” or “extremely poor”. In winter season, the Beel although less affected by surface run-off has a very low water depth thus damping the effects of dilution in the wetland. This may be a plausible reason for elevated concentrations of heavy metal concentration in winter also.

Table 7.2. Entropy weight of physico-chemical parameters

	pH	EC	Turbidity	TH	TA	DO	BOD	TDS	Na^{+}	K^{+}	Mg^{2+}	Ca^{2+}	F^{-}	Cl^{-}	$(SO_4)^{2-}$	$(NO_3)^{-}$
Rainy	0.035	0.056	0.076	0.022	0.031	0.024	0.073	0.024	0.034	0.104	0.049	0.033	0.061	0.134	0.116	0.128
Post-monsoon	0.041	0.061	0.041	0.062	0.061	0.044	0.132	0.059	0.048	0.019	0.062	0.019	0.026	0.082	0.109	0.134
Winter	0.015	0.017	0.062	0.047	0.038	0.059	0.025	0.153	0.143	0.052	0.037	0.032	0.042	0.081	0.074	0.123
Summer	0.019	0.065	0.068	0.041	0.015	0.058	0.092	0.066	0.051	0.097	0.028	0.034	0.023	0.160	0.074	0.110

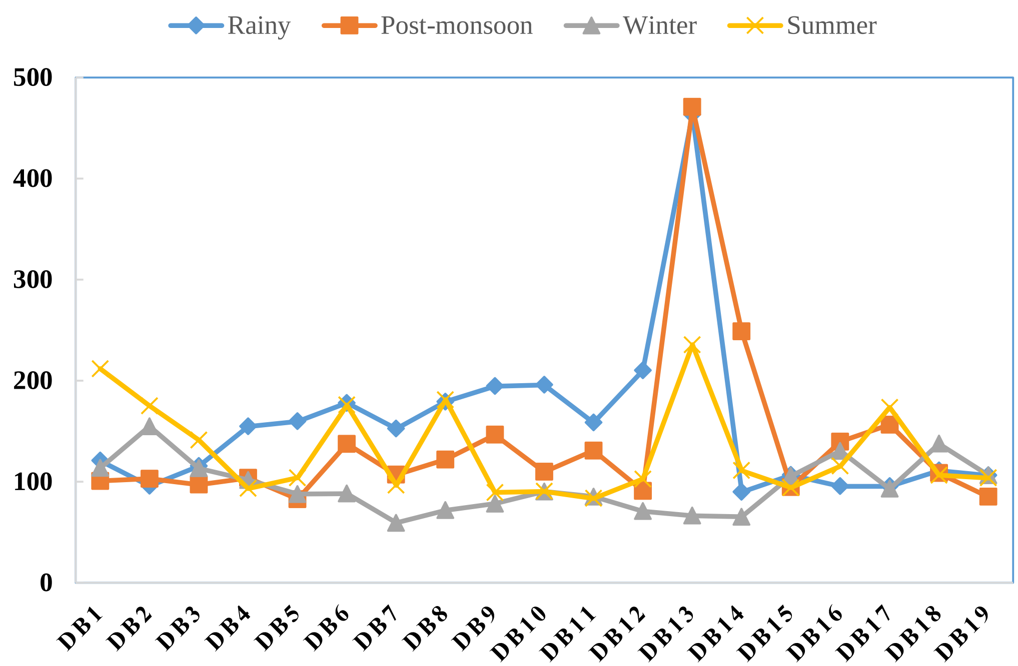


Fig. 7.3. Variation of EWQI in Deepor Beel

Table 7.3 Statistical summary of heavy metals

Sampling Site		Cr	Pb	Fe	Mn	Cu	Zn
DB1	Max	0.16	0.51	4.46	0.47	0.22	0.13
	Min	BDL	BDL	0.70	BDL	0.06	0.02
	Avg	0.07	0.21	2.09	0.18	0.15	0.06
	SD	0.07	0.26	1.64	0.23	0.09	0.05
DB2	Max	0.15	0.35	2.77	0.70	0.25	0.16
	Min	0.01	BDL	0.59	BDL	0.06	0.03
	Avg	0.05	0.16	1.85	0.34	0.16	0.12
	SD	0.06	0.19	0.91	0.39	0.09	0.06
DB3	Max	0.13	0.49	3.23	0.61	0.22	0.20
	Min	0.01	BDL	0.29	BDL	0.06	0.04
	Avg	0.06	0.19	1.46	0.30	0.14	0.10
	SD	0.06	0.24	1.27	0.29	0.09	0.07
DB4	Max	0.11	0.53	1.60	0.31	0.23	0.13
	Min	0.01	BDL	0.26	BDL	0.07	0.01
	Avg	0.04	0.21	0.79	0.12	0.15	0.09
	SD	0.04	0.26	0.57	0.15	0.08	0.06
DB5	Max	0.10	0.54	2.23	0.59	0.25	0.15
	Min	0.01	BDL	0.81	BDL	0.05	BDL
	Avg	0.04	0.25	1.27	0.17	0.19	0.09
	SD	0.04	0.30	0.65	0.28	0.09	0.06
DB6	Max	0.13	0.29	2.82	0.52	0.22	0.13
	Min	0.01	BDL	BDL	BDL	0.07	0.02
	Avg	0.07	0.11	1.15	0.18	0.15	0.08
	SD	0.06	0.14	1.24	0.24	0.08	0.05
DB7	Max	0.14	0.39	1.51	0.31	0.22	0.24
	Min	0.01	BDL	0.72	BDL	0.07	BDL
	Avg	0.06	0.16	1.25	0.11	0.15	0.12
	SD	0.06	0.19	0.37	0.15	0.08	0.10
DB8	Max	0.24	0.29	1.64	0.27	0.24	0.25
	Min	BDL	BDL	0.13	BDL	0.06	0.14
	Avg	0.11	0.14	1.06	0.12	0.16	0.17
	SD	0.11	0.16	0.66	0.14	0.09	0.05
DB9	Max	0.12	0.40	2.81	0.23	0.22	0.17
	Min	0.02	BDL	0.26	BDL	0.06	BDL
	Avg	0.07	0.19	1.46	0.09	0.15	0.09
	SD	0.06	0.22	1.07	0.11	0.08	0.07
DB10	Max	0.09	0.58	2.37	0.33	0.23	0.32
	Min	0.01	BDL	0.71	BDL	0.05	0.05

	Avg	0.03	0.26	1.62	0.19	0.14	0.17
	SD	0.04	0.30	0.79	0.15	0.10	0.11
DB11	Max	0.11	0.43	2.67	0.41	0.23	0.23
	Min	0.01	BDL	0.45	BDL	0.05	BDL
	Avg	0.04	0.21	1.35	0.13	0.14	0.10
	SD	0.05	0.25	0.95	0.19	0.10	0.10
DB12	Max	0.11	0.38	1.54	0.10	0.22	0.16
	Min	BDL	BDL	0.84	BDL	BDL	0.12
	Avg	0.04	0.09	1.20	0.05	0.09	0.13
	SD	0.05	0.19	0.30	0.05	0.09	0.02
DB13	Max	0.12	0.31	6.44	0.45	0.40	0.22
	Min	0.03	BDL	1.29	0.07	0.09	BDL
	Avg	0.05	0.15	3.70	0.31	0.21	0.12
	SD	0.05	0.18	2.11	0.17	0.14	0.09
DB14	Max	0.04	0.46	1.30	0.46	0.22	0.16
	Min	0.02	BDL	0.35	BDL	0.05	BDL
	Avg	0.03	0.22	0.91	0.16	0.14	0.10
	SD	0.01	0.26	0.40	0.22	0.10	0.07
DB15	Max	0.12	0.98	1.69	0.54	0.28	0.26
	Min	0.02	BDL	0.05	BDL	0.10	0.02
	Avg	0.05	0.31	0.85	0.19	0.17	0.12
	SD	0.05	0.46	0.68	0.25	0.09	0.10
DB16	Max	0.07	0.60	3.91	0.47	0.22	0.17
	Min	BDL	BDL	BDL	BDL	BDL	BDL
	Avg	0.02	0.15	1.62	0.21	0.09	0.09
	SD	0.03	0.30	1.75	0.24	0.09	0.09
DB17	Max	0.04	0.24	1.89	0.46	0.26	0.15
	Min	0.01	BDL	0.10	BDL	0.12	0.02
	Avg	0.02	0.10	0.90	0.18	0.20	0.10
	SD	0.01	0.12	0.77	0.24	0.06	0.06
DB18	Max	0.13	0.47	2.77	0.55	0.22	0.13
	Min	0.02	BDL	0.20	BDL	0.04	0.04
	Avg	0.05	0.19	1.17	0.16	0.13	0.11
	SD	0.05	0.23	1.20	0.26	0.10	0.04
DB19	Max	0.09	0.28	2.01	0.19	0.24	0.16
	Min	BDL	BDL	0.16	BDL	0.02	0.02
	Avg	0.03	0.13	1.09	0.06	0.13	0.12
	SD	0.04	0.15	0.78	0.09	0.11	0.07

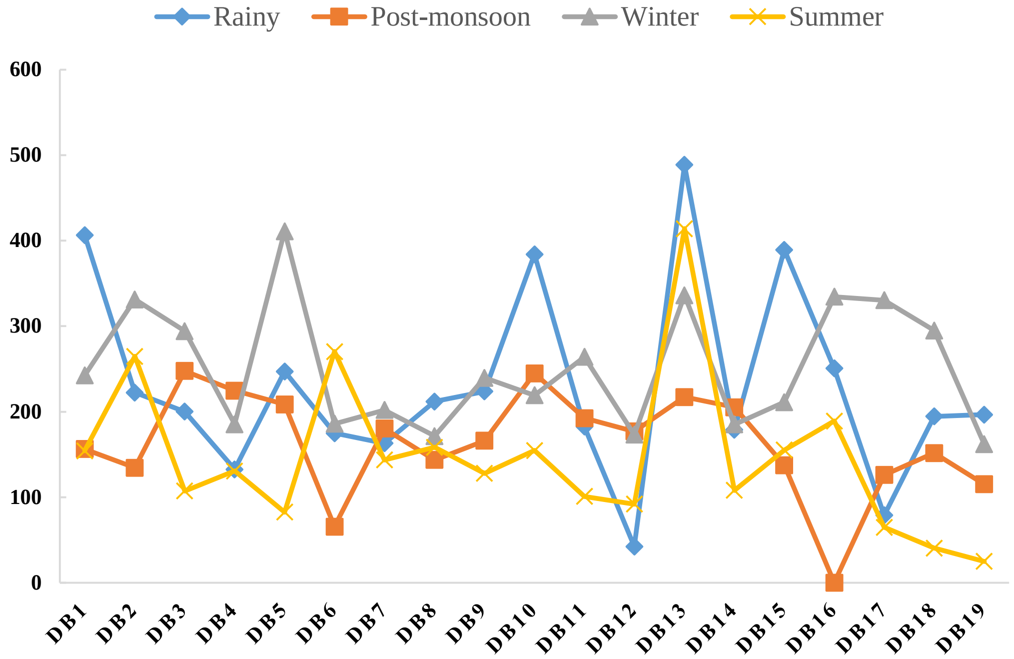


Fig. 7.4. Variation of EHCI in Deepor Beel

7.4 GEOSPATIAL ANALYSIS OF WATER QUALITY VARIABILITY

One-way ANOVA was performed on the EWQI and EHCI values of the different seasons so as to compare the variance between the different groups. The one-way ANOVA test results have been shown in the Table 7.4 and Table 7.5. The test results revealed that there are significant mean differences ($p < 0.05$) between the EWQI values of different seasons as well as the EHCI values of different seasons. This required further investigation of the spatial variability occurring at a sampling station over the period of time with the aid of entropy based disorder indices.

Table 7.4 ANOVA analysis for EWQI

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	39839.03	3	13279.68	2.917929	0.04031	2.739502
Within Groups	309472.2	68	4551.062			
Total	349311.3	71				

Table 7.5 ANOVA analysis for EHCI

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	146975.9	3	48991.97	6.504674	0.000589	2.731807
Within Groups	542290.4	72	7531.811			
Total	689266.3	75				

The study was further carried out using the Shannon's diversity index because entropy is a more accurate measure of dispersion of the probability density function of a data set (Mishra et al., 2009). The geospatial analysis of water quality variability has been shown in the Fig 7.5 (a – d). The geospatial analysis was done using the "Spatial Analyst" tool in the Arc Toolbox of Arc Map 10.2. With reference to the Fig. 7.5 (a), the highest variability of water quality with respect to physico-chemical parameters has been depicted by the EWQI DI. In the vast area of the wetland, relatively higher variability was observed in the three sites namely Site 1, Site 2, and Site 3. In site 1, the sampling stations DB10, DB11, and DB12 depicted high disorder indices contributing to higher variability in the vicinity of these sampling locations over the monitoring period of 2017-2018. Aforementioned influences of dumping, proximity to roadside, and sewage effluents of Guwahati city caused high variability of water quality at these locations. It is noteworthy that the location DB13 had low EWQI DI because it had degraded water quality for majority of the seasons as a consequence of the pronounced effects of leaching at this location. However, the regions with the highest variability in site 1 indicate those that have witnessed highest effects of dilution, surface run-off as well as lowering of water depth throughout the monitoring period. Similarly, in site 2, locations such as DB17, DB18 and DB19 had highest variability of EWQI DI around its vicinity due to probable causes of intermittent discharge of pollution loads from nearby industries and domestic activities over the sampling period. In site 3, high variability was observed due to possible influence of roadside activities and small manufacturing industries set up at National Highway 37 (NH 37) as well as the influence of Khonajan rivulet in the northern side of the wetland. Further investigation of BOD and COD diversity of the Beel yielded similar results as shown in Fig. 7.5 (b) and (c) respectively.

The Fig. 7.5 (d) depicts variability in the four sites of Deepor Beel with respect to EHCI DI. Relatively higher variability of heavy metals was observed in site 2 and site 3 respectively. In site 2, due to presence of settlements and industries, heavy metals may be entering the wetland through intermittent discharge of effluents from the industries over the monitoring period. Also, in site 3 the influences of Khonajan rivulet and a host of manufacturing industries has significant effect on EHCI DI variability. In site 1, the EHCI DI was high for all the seasons due to the accumulation, persistent and non-biodegradable character of heavy metals. This resulted in lower variability of EHCI in this region.

Water quality monitoring networks require optimization of information needs on water quality and the information gathered from monitoring programs (Ozkul et al., 2000). The present entropy-based investigation on the variability of water quality identifies locations which have witnessed the highest influence of cumulative effects of both natural and anthropogenic activities on the water quality. The monitoring of such sites are essential to gather information on the entry and retreat of pollution sources into the wetland, Deepor Beel which has undergone deterioration in recent years due to a host of settlement and resettlement of communities, solid waste dumping, industrial and domestic effluents. The locations DB1, DB10, DB11, DB12 and DB18 require regular monitoring of physico-chemical parameters while the locations DB1, DB4, DB15 and DB17 require regular monitoring of heavy metals.

7.5 RISK ASSESSMENT OF HEAVY METALS ON HUMAN HEALTH

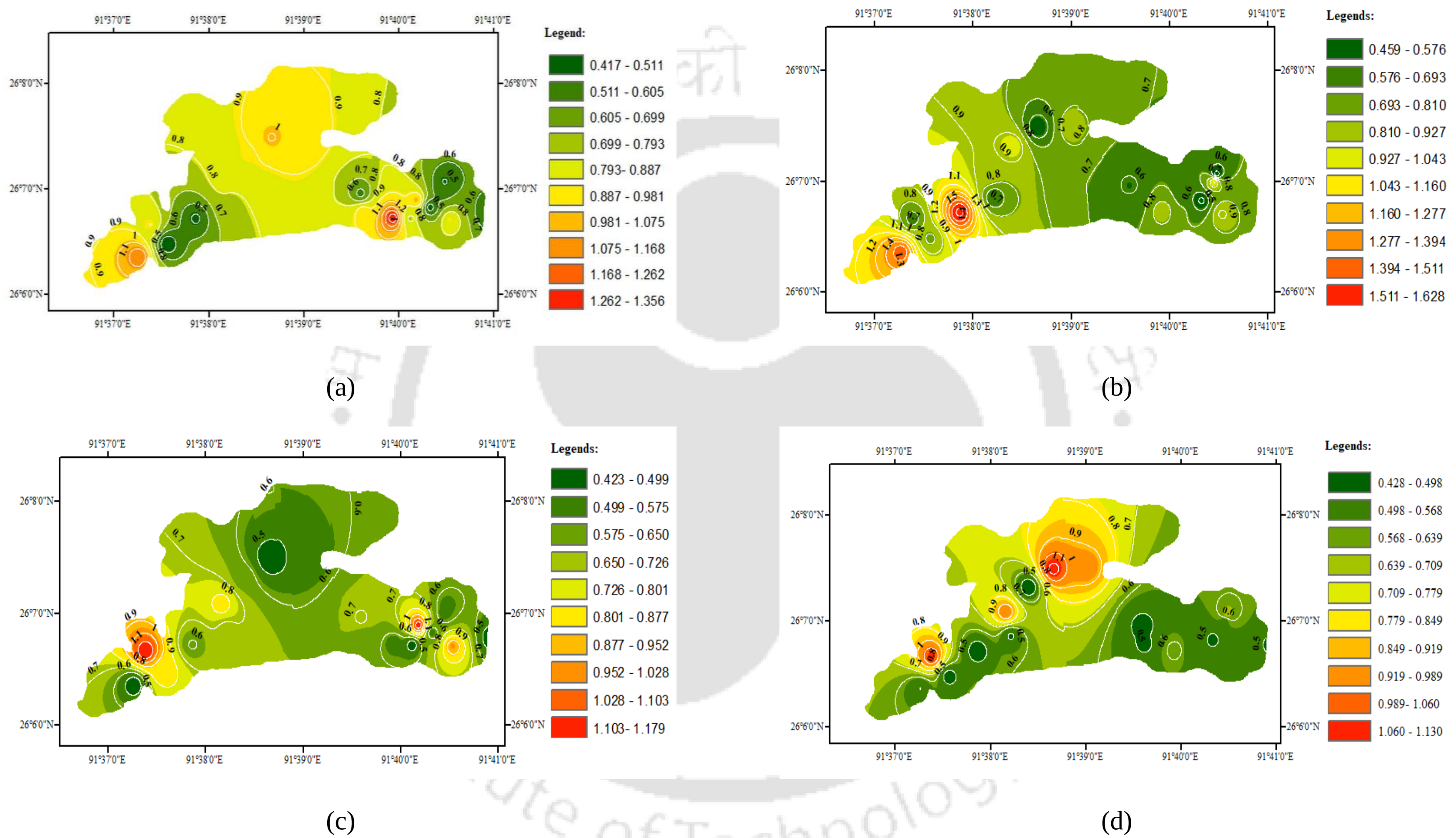
In consideration of the elevated concentrations of heavy metals observed in the wetland in the present study as well as in earlier studies (Kapil and Bhattacharya, 2013; Choudhury and Gupta, 2017), the risk assessment of heavy metals deemed necessary for developing an insight upon the potential threats to communities residing near the wetland. As per the report on visit to Deepor Beel in Assam by the Ministry of Environment and Forests, 2008 unregulated fishing practices are prevalent in most parts of the wetland. The application of risk characterization and assessment of heavy metals suggested by US EPA was extended to this study and revealed some important facts to ponder. The potential risk to human health depicted that the oral exposure was the principal exposure route to intake of heavy metals. Cr and Pb posed highest threat to human health with HI > 1 (1.36 and 9.69 respectively) as shown in Table 7.6. This was evident from the elevated concentrations of these metals found in the wetland.

Table 7.6 ADD and Hazard Index of heavy metals

Metal	ADDing	ADDder	HQi	HQd	HI
Fe	105.29	0.28	0.15	0.00	0.15
Mn	12.72	0.03	0.53	0.04	0.57
Cr	3.69	0.01	1.23	0.13	1.36
Pb	13.56	0.00	9.68	0.01	9.69
Cu	11.17	0.03	0.28	0.00	0.28
Zn	8.18	0.01	0.03	0.00	0.03

7.6 OVERALL RANKING OF SAMPLING SITES

The TOPSIS methodology has been employed to rank the sampling locations of Deepor Beel in each respective season as per their relative pollution levels by taking into account both physico-chemical parameters and heavy metals. The TOPSIS ranks of the sampling locations in Deepor Beel have been shown in the Fig. 7.6 (a-d). The most polluted site was ranked the lowest and the least polluted site was ranked the highest.



Disorder indices and Isoinformation lines of (a) EWQI (b) BOD_5 (c) COD (d) EHCI

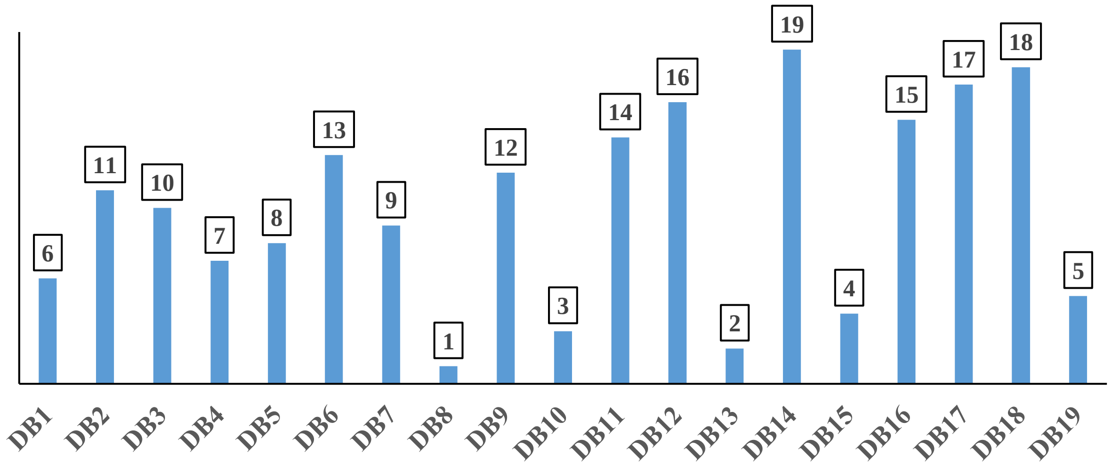


Fig. 7.6 (a). Overall ranking of sampling sites in rainy season

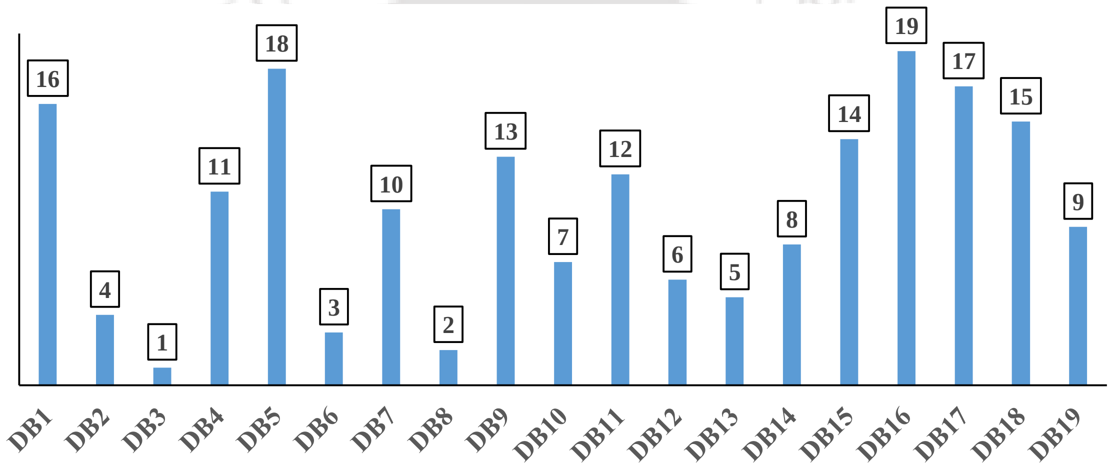


Fig. 7.6 (b). Overall ranking of sampling sites in post monsoon season

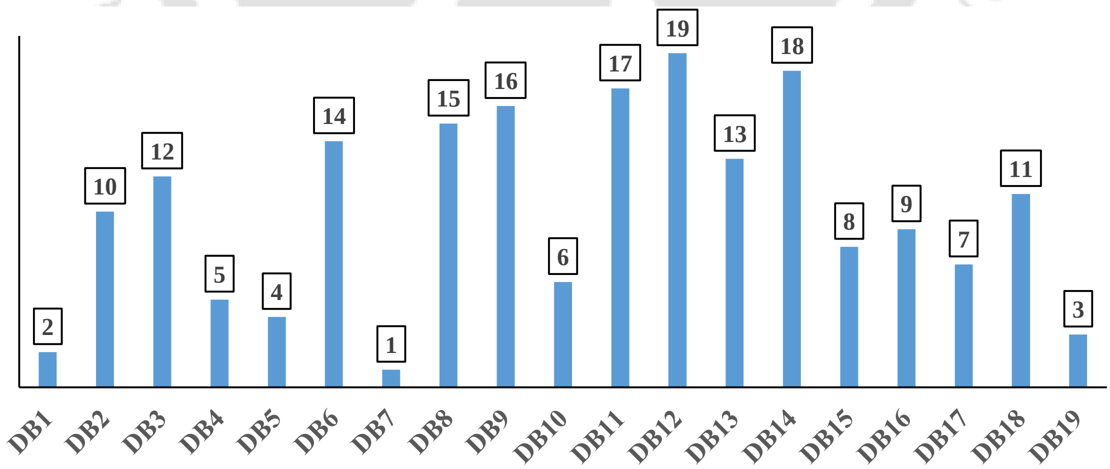


Fig. 7.6 (c). Overall ranking of sampling sites in winter

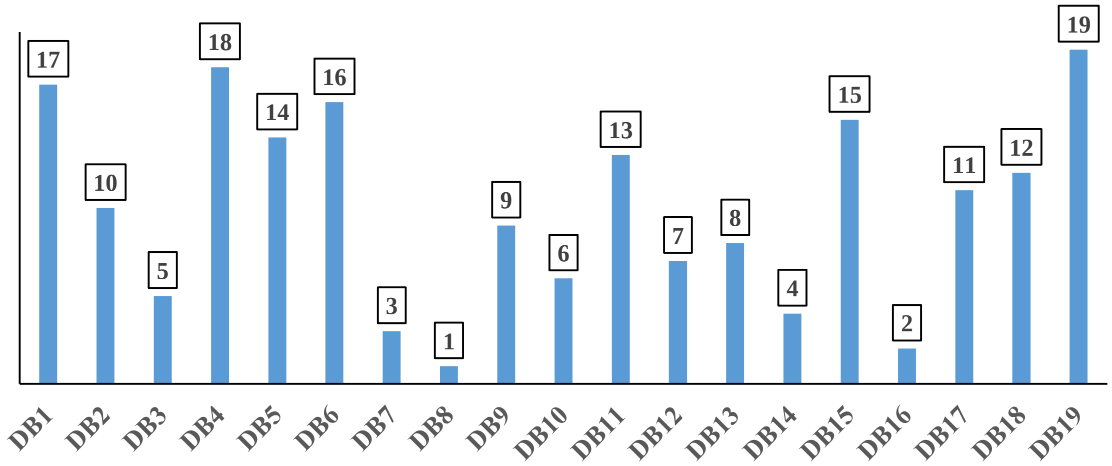


Fig. 7.6 (d). Overall ranking of sampling sites in summer

7.7 CHAPTER CONCLUSION

The important conclusions from the study was drawn as follows:

- The EWQI of the rainy and post-monsoon season were highest at locations of site 1 lying in vicinity of the MSW dumpsite. Such high values owed to the contribution of leaching effects and surface run-off in these locations due to high amount of rainfall.
- Most of peaks of EHCI were observed in the rainy season. The highest EHCI was observed in the DB13 location of site 1. 84.2% of the water samples from the different sampling stations were of "poor" and "very poor" quality. However, all of the sites were contaminated by heavy metals in the winter season which also brought into light the effect of low water depth and lack of dilution in the wetland.
- One-way ANOVA test results on EWQI and EHCI values suggested that there was significant variability of water quality in the wetland in all the four seasons which required further investigation upon quantifying the variability.
- Geospatial analysis of entropy based DI suggested that the locations DB1, DB10, DB11, DB12 and DB18 require regular monitoring of physico-chemical parameters while the locations DB1, DB4, DB15 and DB17 require regular monitoring of heavy metals.

8
Chapter
Conclusions & Recommendations

This chapter covers

- Conclusions of present study
- Recommendations for future work

8.1 CONCLUSIONS

In the present study, some important aspects of water quality monitoring programs have been addressed using environmetrics approach, based on which the following conclusions were summarized:

- The assessment of water quality parameters of Baralia, Puthimari, Pagladia, Beki, Manas, Kameng, Kolong and Brahmaputra Rivers, Assam was achieved with the help of sampling and analysis carried out for the rivers in due course of time. Cluster analysis was applied to group the sampling sites based on their similarity of water quality characteristics. Discriminant analysis and principal component analysis allowed a reduction in the dimensionality of the large datasets indicating few parameters responsible for large variation in water quality. Thus, these statistical tools help in the interpretation of large and complex datasets for water quality assessment,
- An improved water quality index – entropy weighted water quality index (EWQI), was employed subsequently, which takes into account, the uncertainties of occurrences of various physico-chemical parameters. This evaluates the general water quality of any particular water body.
- A similar index - entropy weighted heavy metal contamination index (EHCI) was proposed for heavy metals on the principle of entropy weights to overcome the shortcomings of the existing indices - HPI, CI, HEI and mHPI.

- The use of river water has been further extended for its application in irrigation. For this, an index - entropy weighted irrigation water quality index EIWQI has been generated to determine its applicability in agrarian purposes.
- A reliable entropy based multi-criteria decision making method - TOPSIS has been employed for prioritizing decision makings by ranking the sampling locations on the basis of their overall pollution levels.
- Finally, a novel approach of applying entropy based disorder indices has been adopted to quantify the water quality variability over a temporal scale to assess the cumulative effects of natural and anthropogenic influences. This approach has also been utilized for identifying ideal monitoring locations.

The formulation and use of the present study can thus prove to be of great help to various agencies which take care of the water supply and water pollution control since these form a significant tool for easy understanding and thereby making their applicability uncomplicated. Indeed, these methodologies make the water quality datasets utilization enormously easy and lucid. This study will help policy makers for making decisions in allocating funds and determining priorities. It will also assist in comparing water qualities at different geographical locations, enforcement of water quality standards and determining the changes in water quality. One of the major advantages posed by the methodologies presented in this phase include easy understanding of the outcomes of the analysis, which can thereby, be used to keep the public informed of the overall water quality of any source.

8.2 FUTURE RECOMMENDATIONS

- The present study can be made applicable for other various other sources of water such as ground water, which can give a clearer picture of the ground water table of the particular geographical location.
- Also, several water quality indices can be developed for many other purposes such as industrial, commercial etc., which will take into account their applicability.
- Optimization of water quality monitoring stations can be implemented.
- Other entropy theories such as Rènyi Entropy and Tsallis Entropy may also be investigated for surface water quality assessment.
- Transfer entropy can be used to identify the fate of pollutants.

REFERENCES

Abbasi, T. and Abbasi, S.A., 2012. Water quality indices. Amesterdam Elsevier.

Abdel-Satar, A.M., Ali, M.H. and Goher, M.E., 2017. Indices of water quality and metal pollution of Nile River, Egypt. The Egyptian Journal of Aquatic Research, 43(1), 21-29.

Abtahi, M., Golchinpour, N., Yaghmaeian, K., Rafiee, M., Jahangiri-rad, M., Keyani, A. and Saeedi, R., 2015. A modified drinking water quality index (DWQI) for assessing drinking source water quality in rural communities of Khuzestan Province, Iran. Ecological Indicators, 53, 283-291.

Alobaidy, A.H.M.J., Abid, H.S. and Maulood, B.K., 2010. Application of water quality index for assessment of Dokan lake ecosystem, Kurdistan region, Iraq. Journal of Water Resource and Protection, 2(09), 792.

Alobaidy, A.H.M.J., Al-Sameraiy, M.A., Kadhem, A.J. and Majeed, A.A., 2010. Evaluation of treated municipal wastewater quality for irrigation. Journal of Environmental Protection, 1(03), 216.

Amiri, V., Rezaei, M. and Sohrabi, N., 2014. Groundwater quality assessment using entropy weighted water quality index (EWQI) in Lenjanat, Iran. Environmental Earth Sciences, 72(9), 3479-3490.

Amorocho, J. and Espildora, B., 1973. Entropy in the assessment of uncertainty in hydrologic systems and models. Water Resources Research, 9(6), 1511-1522.

Anderson, A.C., 1994. Iron poisoning in children. Current opinion in pediatrics, 6(3), 289-294.

APHA (2012) Standard Method for the Examination of Water and Wastewater, 22nd edn. American Public Health Association.

Armitage, P.D., Bowes, M.J. and Vincent, H.M., 2007. Long-term changes in macroinvertebrate communities of a heavy metal polluted stream: the river Nent (Cumbria, UK) after 28 years. River Research and Applications, 23(9), 997-1015.

Arveti, N., Sarma, M.R.S., Aitkenhead-Peterson, J.A. and Sunil, K., 2011. Fluoride incidence in groundwater: a case study from Talupula, Andhra Pradesh, India. Environmental Monitoring and Assessment, 172(1-4), 427-443.

Ayers, R.S. and Westcot, D.W., 1985. Water quality for agriculture (Vol. 29). Rome: Food and Agriculture Organization of the United Nations.

Babu, M.T., Das, V.K. and Vethamony, P., 2006. BOD–DO modeling and water quality analysis of a waste water outfall off Kochi, west coast of India. Environment International, 32(2), 165-173.

Backman, B., Bodiš, D., Lahermo, P., Rapant, S. and Tarvainen, T., 1998. Application of a groundwater contamination index in Finland and Slovakia. Environmental Geology, 36(1-2), 55-64.

Basant, N., Gupta, S., Malik, A. and Singh, K.P., 2010. Linear and nonlinear modeling for simultaneous prediction of dissolved oxygen and biochemical oxygen demand of the surface water—a case study. Chemometrics and Intelligent Laboratory Systems, 104(2), 172-180.

Bastami, K.D., Bagheri, H., Kheirabadi, V., Zaferani, G.G., Teymori, M.B., Hamzehpoor, A., Soltani, F., Haghparast, S., Harami, S.R.M., Ghorghani, N.F. and Ganji, S., 2014. Distribution and ecological risk assessment of heavy metals in surface sediments along southeast coast of the Caspian Sea. Marine Pollution Bulletin, 81(1), 262-267.

Batayneh, A. and Zumlot, T., 2012. Multivariate statistical approach to geochemical methods in water quality factor identification; application to the shallow aquifer system of the Yarmouk Basin of north Jordan. Research Journal of Environmental and Earth Sciences, 4(7), 756-768.

Beamonte, E., Bermúdez, J.D., Casino, A. and Veres, E., 2005. A global stochastic index for water quality: the case of the river Turia in Spain. Journal of Agricultural, Biological, and Environmental Statistics, 10(4), 424.

Behmel, S., Damour, M., Ludwig, R. and Rodriguez, M.J., 2016. Water quality monitoring strategies—a review and future perspectives. Science of the Total Environment, 571, 1312-1329.

Bhargava, D.S., 1983. Use of water quality index for river classification and zoning of Ganga River. Environmental Pollution Series B, Chemical and Physical, 6(1), 51-67.

Bhuiyan, M.A., Islam, M.A., Dampare, S.B., Parvez, L. and Suzuki, S., 2010. Evaluation of hazardous metal pollution in irrigation and drinking water systems in the vicinity of a coal mine area of northwestern Bangladesh. Journal of Hazardous Materials, 179(1-3), 1065-1077.

Bhuiyan, M.A.H., Dampare, S.B., Islam, M.A. and Suzuki, S., 2015. Source apportionment and pollution evaluation of heavy metals in water and sediments of Buriganga River, Bangladesh, using multivariate analysis and pollution evaluation indices. Environmental Monitoring and Assessment, 187(1), 4075.

Bhuyan, M.S., Bakar, M.A., Akhtar, A., Hossain, M.B., Ali, M.M. and Islam, M.S., 2017. Heavy metal contamination in surface water and sediment of the Meghna River, Bangladesh. Environmental Nanotechnology, Monitoring & Management, 8, 273-279.

BIS IS10500, 2012. Indian Standard Drinking Water–Specification (Second revision). Bureau of Indian Standards (BIS), New Delhi.

Bouguerne, A., Boudoukha, A., Benkhaled, A. and Mebarkia, A.H., 2017. Assessment of surface water quality of Ain Zada dam (Algeria) using multivariate statistical techniques. International Journal of River Basin Management, 15(2), 133-143.

Boyacioglu, H. and Boyacioglu, H., 2008. Water pollution sources assessment by multivariate statistical methods in the Tahtali Basin, Turkey. Environmental Geology, 54(2), 275-282.

Brown, R.M., McClelland, N.I., Deininger, R.A. and Tozer, R.G., 1970. A Water Quality Index-Do We Dare.

Brunsell, N.A., 2010. A multiscale information theory approach to assess spatial–temporal variability of daily precipitation. Journal of Hydrology, 385(1-4), 165-172.

Bu, H., Tan, X., Li, S. and Zhang, Q., 2010. Water quality assessment of the Jinshui River (China) using multivariate statistical techniques. Environmental Earth Sciences, 60(8), 1631-1639.

Carpenter, S.R., Caraco, N.F., Correll, D.L., Howarth, R.W., Sharpley, A.N. and Smith, V.H., 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecological Applications, 8(3), 559-568.

Chaturvedi, A., Bhattacharjee, S., Singh, A.K. and Kumar, V., 2018. A new approach for indexing groundwater heavy metal pollution. Ecological Indicators, 87, 323-331.

Cheng, W.L., Kuo, Y.C., Lin, P.L., Chang, K.H., Chen, Y.S., Lin, T.M. and Huang, R., 2004. Revised air quality index derived from an entropy function. Atmospheric Environment, 38(3), 383-391

Choudhury, D. and Gupta, S., 2017. Impact of waste dump on surface water quality and aquatic insect diversity of Deepor Beel (Ramsar site), Assam, North-east India. Environmental Monitoring and Assessment, 189(11), 540.

Christian, R.R. and Pipes, W.O., 1983. Frequency distribution of coliforms in water distribution systems. Applied and Environmental Microbiology, 45(2), 603-609.

Chung, S.Y., Venkatramanan, S., Kim, T.H., Kim, D.S. and Ramkumar, T., 2015. Influence of hydrogeochemical processes and assessment of suitability for groundwater uses in Busan City, Korea. Environment, Development and Sustainability, 17(3), 423-441.

Cox, B.A., 2003. A review of dissolved oxygen modelling techniques for lowland rivers. Science of the Total Environment, 314, 303-334.

Cude, C.G., 2001. Oregon Water Quality Index a Tool for Evaluating Water Quality Management Effectiveness 1. JAWRA Journal of the American Water Resources Association, 37(1), 125-137.

Davraz, A. and Özdemir, A., 2014. Groundwater quality assessment and its suitability in Çeltikçi plain (Burdur/Turkey). Environmental Earth Sciences, 72(4), 1167-1190.

Dixon, W. and Chiswell, B., 1996. Review of aquatic monitoring program design. Water research, 30(9), 1935-1948.

Djebou, D.C.S., Singh, V.P. and Frauenfeld, O.W., 2014. Analysis of watershed topography effects on summer precipitation variability in the southwestern United States. Journal of Hydrology, 511, 838-849.

Dojlido, J., Raniszewski, J. and Woyciechowska, J., 1994. Water quality index-application for rivers in Vistula river basin in Poland. Water Science and Technology, 30(10), 57.

Edet, A.E. and Offiong, O.E., 2002. Evaluation of water quality pollution indices for heavy metal contamination monitoring. A study case from Akpabuyo-Odukpani area, Lower Cross River Basin (southeastern Nigeria). GeoJournal, 57(4), 295-304.

Einax, J.W., Truckenbrodt, D. and Kampe, O., 1998. River pollution data interpreted by means of chemometric methods. Microchemical journal, 58(3), 315-324.

Fagbote, E.O., Olanipekun, E.O. and Uyi, H.S., 2014. Water quality index of the ground water of bitumen deposit impacted farm settlements using entropy weighted method. International Journal of Environmental Science and Technology, 11(1), 127-138.

Faiz, M.A., Liu, D., Fu, Q., Wrzesiński, D., Muneer, S., Khan, M.I., Li, T. and Cui, S., 2017. Assessment of precipitation variability and uncertainty of stream flow in the Hindu Kush Himalayan and Karakoram River basins of Pakistan. Meteorology and Atmospheric Physics, 1-10.

Falah, F. and Haghizadeh, A., 2017. Hydrochemical evaluation of river water quality—a case study: Horroud River. Applied Water Science, 7(8), 4725-4733.

Fan, X., Cui, B., Zhao, H., Zhang, Z. and Zhang, H., 2010. Assessment of river water quality in Pearl River Delta using multivariate statistical techniques. Procedia Environmental Sciences, 2, 1220-1234.

Fernández, N., Ramírez, A. and Solano, F., 2004. Physico-chemical water quality indices-A comparative review–. Bistua Revista de la Facultad de Ciencias Basicas, 2(1).

Field, A. (2000). Discovering statistics using SPSS for Windows. London – Thousand Oaks – New Delhi: Sage publication.

Fleming, S.W., 2007. An information theoretic perspective on mesoscale seasonal variations in ground-level ozone. Atmospheric Environment, 41(27), 5746-5755

Fosmire, G.J., 1990. Zinc toxicity. The American journal of clinical nutrition, 51(2), 225-227.

Fu, J., Zhao, C., Luo, Y., Liu, C., Kyzas, G.Z., Luo, Y., Zhao, D., An, S. and Zhu, H., 2014. Heavy metals in surface sediments of the Jialu River, China: their relations to environmental factors. Journal of Hazardous Materials, 270, 102-109.

Giri, S. and Singh, A.K., 2014. Assessment of surface water quality using heavy metal pollution index in Subarnarekha River, India. Water Quality, Exposure and Health, 5(4), 173-182.

Giri, S., Singh, G., Gupta, S.K., Jha, V.N. and Tripathi, R.M., 2010. An evaluation of metal contamination in surface and groundwater around a proposed uranium mining site, Jharkhand, India. Mine Water and the Environment, 29(3), 225-234.

Goldhaber, S.B., 2003. Trace element risk assessment: essentiality vs. toxicity. Regulatory Toxicology and Pharmacology, 38(2), 232-242.

Gorgij, A.D., Kisi, O., Moghaddam, A.A. and Taghipour, A., 2017. Groundwater quality ranking for drinking purposes, using the entropy method and the spatial autocorrelation index. Environmental Earth Sciences, 76(7), 269.

Gupta, S.K., Gupta, R.C., Gupta, A.B., Seth, A.K., Bassin, J.K., Gupta, A. and Sharma, M.L., 2001. Recurrent diarrhea in children living in areas with high levels of nitrate in drinking water. Archives of Environmental Health: An International Journal, 56(4), 369-373.

Hajigholizadeh, M. and Melesse, A.M., 2017. Assortment and spatiotemporal analysis of surface water quality using cluster and discriminant analyses. Catena, 151, 247-258.

Harkins, R.D., 1974. An objective water quality index. Journal (Water Pollution Control Federation), 588-591.

Harmanescu, M., Alda, L.M., Bordean, D.M., Gogoasa, I. and Gergen, I., 2011. Heavy metals health risk assessment for population via consumption of vegetables grown in old mining area; a case study: Banat County, Romania. Chemistry Central Journal, 5(1), 64.

Helena, B., Pardo, R., Vega, M., Barrado, E., Fernandez, J.M. and Fernandez, L., 2000. Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis. Water Research, 34(3), 807-816.

Hernandez, E.A. and Uddameri, V., 2010. Selecting agricultural best management practices for water conservation and quality improvements using Atanassov's intuitionistic fuzzy sets. Water Resources Management, 24(15), 4589-4612.

Horton, R.K., 1965. An index number system for rating water quality. Journal of Water Pollution Control Federation, 37(3), 300-306.

Hsu, H.M., Moncrieff, M.W., Tung, W.W. and Liu, C., 2006. Multiscale temporal variability of warm-season precipitation over North America: Statistical analysis of radar measurements. Journal of the Atmospheric Sciences, 63(9), 2355-2368.

Hwang, C.L., Lai, Y.J. and Liu, T.Y., 1993. A new approach for multiple objective decision making. Computers & Operations Research, 20(8), 889-899.

Iscen, C.F., Emiroglu, Ö., Ilhan, S., Arslan, N., Yilmaz, V. and 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.

Islam, A.R.M.T., Ahmed, N., Bodrud-Doza, M. and Chu, R., 2017. Characterizing groundwater quality ranks for drinking purposes in Sylhet district, Bangladesh, using entropy method, spatial autocorrelation index, and geostatistics. Environmental Science and Pollution Research, 24(34), 26350-26374.

Islam, M.S., Ahmed, M.K., Raknuzzaman, M., Habibullah-Al-Mamun, M. and Islam, M.K., 2015. Heavy metal pollution in surface water and sediment: a preliminary assessment of an urban river in a developing country. Ecological Indicators, 48, 282-291.

Jeong, Y., Sanders, B.F. and Grant, S.B., 2006. The information content of high-frequency environmental monitoring data signals pollution events in the coastal ocean. Environmental Science & Technology, 40(20), 6215-6220.

Jha, R., Ojha, C.S. and Bhatia, K.K., 2007. Development of refined BOD and DO models for highly polluted Kali River in India. Journal of Environmental Engineering, 133(8), 839-852.

Jianhua, W., Peiyue, L. and Hui, Q., 2011. Groundwater quality in Jingyuan County, a semi-humid area in Northwest China. Journal of Chemistry, 8(2), 787-793.

Jianqin, M., Jingjing, G. and Xiaojie, L., 2010. Water quality evaluation model based on principal component analysis and information entropy: application in Jinshui River. Journal of Resources and Ecology, 1(3), 249-252.

Jung, K.Y., Lee, K.L., Im, T.H., Lee, I.J., Kim, S., Han, K.Y. and Ahn, J.M., 2016. Evaluation of water quality for the Nakdong River watershed using multivariate analysis. Environmental Technology & Innovation, 5, 67-82.

Kaiser, H.F., 1960. The application of electronic computers to factor analysis. Educational and Psychological Measurement, 20(1), 141-151

Kalita, G.J., Sarma, P.K., Goswami, P. and Rout, S., 2015. Socio-economic status of fishermen and different fishing gear used in Beki River, Barpeta, Assam. Journal of Entomology and Zoology Studies, 3(1), 193-198.

Kapil, N. and Bhattacharyya, K.G., 2012. Spatial, temporal and depth profiles of trace metals in an urban Wetland system: A case study with respect to the deepor beel, Ramsar site 1207, India. Environment and Pollution, 2(1), 51.

Karak, T., Bhattacharyya, P., Das, T., Paul, R.K. and Bezbaruah, R., 2013. Non-segregated municipal solid waste in an open dumping ground: a potential contaminant in relation to environmental health. International Journal of Environmental Science and Technology, 10(3), 503-518.

Karamouz, M., Nokhandan, A.K., Kerachian, R. and Maksimovic, Č., 2009. Design of on-line river water quality monitoring systems using the entropy theory: a case study. Environmental Monitoring and Assessment, 155(1), 63-81.

Kaurish, F.W. and Younos, T., 2007. Developing a standardized water quality index for evaluating surface water quality. JAWRA Journal of the American Water Resources Association, 43(2), 533-545.

Kazi, T.G., Arain, M.B., Jamali, M.K., Jalbani, N., Afridi, H.I., Sarfraz, R.A., Baig, J.A. and 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.

Keum, J., Kornelsen, K.C., Leach, J.M. and Coulibaly, P., 2017. Entropy Applications to Water Monitoring Network Design: A Review. Entropy, 19(11), 613.

Khan, A. and Maity, K., 2016. Application of MCDM-based TOPSIS Method for the Optimization of Multi Quality Characteristics of Modern Manufacturing Processes. International Journal of Engineering Research in Africa, 23.

Khan, M.U., Malik, R.N. and Muhammad, S., 2013. Human health risk from heavy metal via food crops consumption with wastewater irrigation practices in Pakistan. Chemosphere, 93(10), 2230-2238.

Khodapanah, L.W.N.A., Sulaiman, W.N.A. and Khodapanah, N., 2009. Groundwater quality assessment for different purposes in Eshtehard District, Tehran, Iran. European Journal of Scientific Research, 36(4), 543-553.

Khound, N.J. and Bhattacharyya, K.G., 2017. Multivariate statistical evaluation of heavy metals in the surface water sources of Jia Bharali river basin, North Brahmaputra plain, India. Applied Water Science, 7(5), 2577-2586.

Kim, J. O., Mueller, C. W. 1987. Introduction to factor analysis: what it is and how to do it. Quantitative Applications in the Social Sciences Series, Sage University Press, Newbury Park.

Konstantinou, I.K., Hela, D.G. and Albanis, T.A., 2006. The status of pesticide pollution in surface waters (rivers and lakes) of Greece. Part I. Review on occurrence and levels. Environmental Pollution, 141(3), 555-570.

Krishna, A.K., Satyanarayanan, M. and Govil, P.K., 2009. Assessment of heavy metal pollution in water using multivariate statistical techniques in an industrial area: a case study from Patancheru, Medak District, Andhra Pradesh, India. Journal of Hazardous Materials, 167(1), 366-373.

Kumar, P.S., Delson, P.D. and Babu, P.T., 2012. Appraisal of heavy metals in groundwater in Chennai city using a HPI model. Bulletin of Environmental Contamination and Toxicology, 89(4), 793-798.

Lamson, D.W. and Plaza, S.M., 2002. The safety and efficacy of high-dose chromium.(High-Dose Chromium). Alternative Medicine Review, 7(3), 218-236.

Lande, R., 1996. Statistics and partitioning of species diversity, and similarity among multiple communities. Oikos, 5-13.

Landwehr, J.M., 1979. A statistical view of a class of water quality indices. Water Resources Research, 15(2), 460-468.

Li, P., Li, X., Meng, X., Li, M. and Zhang, Y., 2016. Appraising groundwater quality and health risks from contamination in a semiarid region of northwest China. Exposure and Health, 8(3), 361-379.

Li, R., Zou, Z. and An, Y., 2016. Water quality assessment in Qu River based on fuzzy water pollution index method. Journal of Environmental Sciences, 50, 87-92.

Li, X., Wang, K., Liu, L., Xin, J., Yang, H. and Gao, C., 2011. Application of the entropy weight and TOPSIS method in safety evaluation of coal mines. Procedia Engineering, 26, 2085-2091.

Ling, T.Y., Soo, C.L., Liew, J.J., Nyanti, L., Sim, S.F. and Grinang, J., 2017. Application of multivariate statistical analysis in evaluation of surface river water quality of a tropical river. Journal of Chemistry, 2017.

Liou, S.M., Lo, S.L. and Wang, S.H., 2004. A generalized water quality index for Taiwan. Environmental Monitoring and Assessment, 96(1-3), 35-52.

Liu, B., Chen, X., Lian, Y. and Wu, L., 2013. Entropy-based assessment and zoning of rainfall distribution. Journal of Hydrology, 490, 32-40.

Liu, L., Zhou, J., An, X., Zhang, Y. and Yang, L., 2010. Using fuzzy theory and information entropy for water quality assessment in Three Gorges region, China. Expert Systems with Applications, 37(3), 2517-2521.

Liu, T.K., Yu, J.L., Chen, C.L. and Wei, P.S., 2012. Information theoretic perspective on coastal water-quality monitoring and management near an offshore industrial park. Environmental Monitoring and Assessment, 184(8), 4725-4735.

Lobato, T.C., Hauser-Davis, R.A., Oliveira, T.F., Silveira, A.M., Silva, H.A.N., Tavares, M.R.M. and Saraiva, A.C.F., 2015. 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.

Lohani, B.N. and Mustapha, N., 1982. Stochastic water quality index. Environmental Technology, 3(1-11), 521-528.

Luo, K., Hu, X., He, Q., Wu, Z., Cheng, H., Hu, Z. and Mazumder, A., 2017. Using multivariate techniques to assess the effects of urbanization on surface water quality: a case study in the Liangjiang New Area, China. Environmental Monitoring and Assessment, 189(4), p.174.

Malik, R.N. and Hashmi, M.Z., 2017. Multivariate statistical techniques for the evaluation of surface water quality of the Himalayan foothills streams, Pakistan. Applied Water Science, 7(6), 2817-2830.

Maltby, E., 2013. Waterlogged wealth: why waste the world's wet places?. Routledge.

Martino, G.D., Fontana, N., Marini, G. and Singh, V.P., 2012. Variability and trend in seasonal precipitation in the continental United States. Journal of Hydrologic Engineering, 18(6), 630-640.

Matta, G., Srivastava, S., Pandey, R.R. and Saini, K.K., 2017. Assessment of physicochemical characteristics of Ganga Canal water quality in Uttarakhand. Environment, Development and Sustainability, 19(2), 419-431.

Maurer, B.A. and McGill, B.J., 2011. Measurement of species diversity. Biological diversity: frontiers in measurement and assessment. Oxford University Press, Oxford, 55-65.

McIntosh, R.P., 1967. An index of diversity and the relation of certain concepts to diversity. Ecology, 48(3), 392-404.

McKenna, J.E., 2003. An enhanced cluster analysis program with bootstrap significance testing for ecological community analysis. Environmental Modelling & Software, 18(3), 205-220.

Misaghi, F., Delgosha, F., Razzaghmanesh, M. and Myers, B., 2017. Introducing a water quality index for assessing water for irrigation purposes: A case study of the Ghezel Ozan River. Science of the Total Environment, 589, 107-116.

Mishra, A.K., Özger, M. and Singh, V.P., 2009. An entropy-based investigation into the variability of precipitation. Journal of Hydrology, 370(1), 139-154.

Mogheir, Y., De Lima, J.L.M.P. and Singh, V.P., 2004. Characterizing the spatial variability of groundwater quality using the entropy theory: I. Synthetic data. Hydrological Processes, 18(11), 2165-2179.

Mohammadpour, R., Shaharuddin, S., Chang, C.K., Zakaria, N.A., Ab Ghani, A. and Chan, N.W., 2015. Prediction of water quality index in constructed wetlands using support vector machine. Environmental Science and Pollution Research, 22(8), 6208-6219.

Mohan, S.V., Nithila, P. and Reddy, S.J., 1996. Estimation of heavy metals in drinking water and development of heavy metal pollution index. Journal of Environmental Science & Health Part A, 31(2), 283-289.

Moulton, I.I., Stephen, R., Carter, J.L., Grotheer, S.A., Cuffney, T.F. and Short, T.M., 2000. Methods of analysis by the US Geological Survey National Water Quality Laboratory-processing, taxonomy, and quality control of benthic macroinvertebrate samples (No. USGS-00-212). Department of the Interior Washington DC.

Nagarajah, S., Abeykoon, V., Emerson, B.N. and Yogalingam, S., 1988. Water quality of some wells in Jaffna and Kilinochchi with special reference to nitrate pollution [Sri Lanka]. Tropical Agriculturist (Sri Lanka).

Noori, R., Sabahi, M.S., Karbassi, A.R., Baghvand, A. and Zadeh, H.T., 2010. Multivariate statistical analysis of surface water quality based on correlations and variations in the data set. Desalination, 260(1), 129-136.

Ocampo-Duque, W., Ferre-Huguet, N., Domingo, J.L. and Schuhmacher, M., 2006. Assessing water quality in rivers with fuzzy inference systems: A case study. Environment International, 32(6), 733-742.

Ongley, E.D., 1998. Modernisation of water quality programmes in developing countries: issues of relevancy and cost efficiency. Water Quality International, 37-42.

Ouyang, Y., Nkedi-Kizza, P., Wu, Q.T., Shinde, D. and Huang, C.H., 2006. Assessment of seasonal variations in surface water quality. Water Research, 40(20), 3800-3810.

Ozkul, S., Harmancioglu, N.B. and Singh, V.P., 2000. Entropy-based assessment of water quality monitoring networks. Journal of Hydrologic Engineering, 5(1), 90-100.

Papanikolaou, N.C., Hatzidaki, E.G., Belivanis, S., Tzanakakis, G.N. and Tsatsakis, A.M., 2005. Lead toxicity update. A brief review. Medical Science Monitor, 11(10), RA329-RA336.

Pei-Yue, L., Hui, Q. and Jian-Hua, W.U., 2010. Groundwater quality assessment based on improved water quality index in Pengyang County, Ningxia, Northwest China. Journal of Chemistry, 7(S1), S209-S216.

Pekey, H., Karakaş, D. and Bakoglu, M., 2004. Source apportionment of trace metals in surface waters of a polluted stream using multivariate statistical analyses. Marine Pollution Bulletin, 49(9-10), 809-818.

Peng, L., 2004. A universal index formula suitable to multiparameter water quality evaluation. Numerical Methods for Partial Differential Equations: An International Journal, 20(3), 368-373.

Pertsemli, E. and Voutsa, D., 2007. Distribution of heavy metals in lakes Doirani and Kerkini, Northern Greece. Journal of Hazardous Materials, 148(3), 529-537.

Prasad, A.S., 1976. Deficiency of Zinc in man and its toxicity. Trace Elements in Health and Disease Vol I, Zinc and Copper.

Prasad, B. and Bose, J., 2001. Evaluation of the heavy metal pollution index for surface and post-monsoon water near a limestone mining area of the lower Himalayas. Environmental Geology, 41(1), 183-188.

Prasad, B. and Sangita, K., 2008. Heavy metal pollution index of ground water of an abandoned open cast mine filled with fly ash: a case study. Mine Water and the Environment, 27(4), 265-267.

Prasad, B., Kumari, P., Bano, S. and Kumari, S., 2014. Ground water quality evaluation near mining area and development of heavy metal pollution index. Applied Water Science, 4(1), 11-17.

Prasanna, M.V., Praveena, S.M., Chidambaram, S., Nagarajan, R. and Elayaraja, A., 2012. Evaluation of water quality pollution indices for heavy metal contamination monitoring: a case study from Curtin Lake, Miri City, East Malaysia. Environmental Earth Sciences, 67(7), 1987-2001.

Prati, L., Pavanello, R. and Pesarin, F., 1971. Assessment of surface water quality by a single index of pollution. Water Research, 5(9), 741-751.

Rahman, M.S., Saha, N., Islam, A.T., Shen, S. and Bodrud-Doza, M., 2017. Evaluation of Water Quality for Sustainable Agriculture in Bangladesh. Water, Air, & Soil Pollution, 228(10), 385.

Rakotondrabe, F., Ngoupayou, J.R.N., Mfonka, Z., Rasolomanana, E.H., Abolo, A.J.N. and Ako, A.A., 2018. Water quality assessment in the Betare-Oya gold mining area (East-Cameroon): multivariate statistical analysis approach. Science of the Total Environment, 610, 831-844.

Ribbe, L., Delgado, P., Salgado, E. and Flügel, W.A., 2008. Nitrate pollution of surface water induced by agricultural non-point pollution in the Pocochay watershed, Chile. Desalination, 226(1-3), 13-20.

Roushangar, K. and Alizadeh, F., 2018. Identifying complexity of annual precipitation variation in Iran during 1960–2010 based on information theory and discrete wavelet transform. Stochastic Environmental Research and Risk Assessment, 32(5), 1205-1223.

Sahoo, M.M. and Patra, K.C., 2018. Development of a Functional Water Quality Index: a case study in Brahmani River Basin, India. International Journal of Environmental Science and Technology, 1-14.

Sahoo, M.M., Patra, K.C., Swain, J.B. and Khatua, K.K., 2017. Evaluation of water quality with application of Bayes' rule and entropy weight method. European Journal of Environmental and Civil Engineering, 21(6), 730-752.

Salifu, M., Aidoo, F., Hayford, M.S., Adomako, D. and Asare, E., 2017. Evaluating the suitability of groundwater for irrigational purposes in some selected districts of the Upper West region of Ghana. Applied Water Science, 7(2), 653-662.

Sargaonkar, A. and Deshpande, V., 2003. Development of an overall index of pollution for surface water based on a general classification scheme in Indian context. Environmental Monitoring and Assessment, 89(1), 43-67.

Sawyer, C.N. and McCarty, P.L., 1978. Chemistry for environmental engineers. New York. Mc Graw-Hill Book Company.

Shammi, M., Rahman, R., Rahman, M.M., Moniruzzaman, M., Bodrud-Doza, M., Karmakar, B. and Uddin, M.K., 2016. Assessment of salinity hazard in existing water resources for irrigation and potentiality of conjunctive uses: a case report from Gopalganj District, Bangladesh. Sustainable Water Resources Management, 2(4), 369-378.

Shannon, C.E., Weaver, W. and Burks, A.W., 1951. The mathematical theory of communication.

Sharma, D.A., Rishi, M.S. and Keesari, T., 2017. Evaluation of groundwater quality and suitability for irrigation and drinking purposes in southwest Punjab, India using hydrochemical approach. Applied Water Science, 7(6), 3137-3150.

Sharma, M.P., Sharma, S., Goel, V., Sharma, P. and Kumar, A., 2006. Water quality assessment of Behta River using benthic macroinvertebrates. Life Science Journal, 3(4), 68-74.

Shihab, A.S. and Abdul Baqi, Y.T., 2010. Multivariate analysis of ground water quality of Makhmor plain/north Iraq. Damascus University Journal, 26(1), 19-26.

Shrestha, S. and 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.

Sianaki, O.A. and Masoum, M.A., 2013. A multi-agent intelligent decision making support system for home energy management in smart grid: A fuzzy TOPSIS approach. Multiagent and Grid Systems, 9(3), 181-195.

Silva, V.D.P.R., Belo Filho, A.F., Almeida, R.S.R., de Holanda, R.M. and da Cunha Campos, J.H.B., 2016. Shannon information entropy for assessing space–time variability of rainfall and streamflow in semiarid region. Science of the Total Environment, 544, 330-338.

Simeonov, V., Stratis, J.A., Samara, C., Zachariadis, G., Voutsa, D., Anthemidis, A., Sofoniou, M. and Kouimtzis, T., 2003. Assessment of the surface water quality in Northern Greece. Water research, 37(17), 4119-4124.

Sin, S.N., Chua, H., Lo, W. and Ng, L.M., 2001. Assessment of heavy metal cations in sediments of Shing Mun River, Hong Kong. Environment International, 26(5), 297-301.

Singh R, Venkatesh AS, Syed TH, Reddy AG, Kumar M, Kurakalva RM. Assessment of potentially toxic trace elements contamination in groundwater resources of the coal mining area of the Korba Coalfield, Central India. Environmental Earth Sciences. 2017 Aug 1;76(16):566.

Singh, H., Singh, D., Singh, S.K. and Shukla, D.N., 2017. Assessment of river water quality and ecological diversity through multivariate statistical techniques, and earth observation dataset of rivers Ghaghara and Gandak, India. International Journal of River Basin Management, 15(3), 347-360.

Singh, K.P., Malik, A. and Sinha, S., 2005. Water quality assessment and apportionment of pollution sources of Gomti river (India) using multivariate statistical techniques—a case study. Analytica Chimica Acta, 538(1-2), 355-374.

Singh, K.P., Malik, A., Mohan, D. and 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.

Singh, V.P., 1997. The use of entropy in hydrology and water resources. Hydrological Processes, 11(6), 587-626.

Singh, V.P., 2000. The entropy theory as a tool for modeling and decision-making in environmental and water resources.

Singh, V.P., 2011. Hydrologic synthesis using entropy theory. Journal of Hydrologic Engineering, 16(5), 421-433.

Singh, V.P., 2013. Entropy theory and its application in environmental and water engineering. John Wiley & Sons.

Singh, V.P., 2014, June. Entropy theory in hydraulic engineering: An introduction. American Society of Civil Engineers.

Sivakumar, B., 2001. Rainfall dynamics at different temporal scales: a chaotic perspective. Hydrology and Earth System Sciences, 5(4), 645-652.

Smith, R.A., Schwarz, G.E. and Alexander, R.B., 1997. Regional interpretation of water-quality monitoring data. Water Resources Research, 33(12), 2781-2798.

Sojka, M., Siepak, M., Zioła, A., Frankowski, M., Murat-Błażejewska, S. and Siepak, J., 2008. Application of multivariate statistical techniques to evaluation of water quality in the Mała Wełna River (Western Poland). Environmental Monitoring and Assessment, 147(1), 159-170.

Soofi, E.S., 1997. Information theoretic regression methods. In Applying Maximum Entropy to Econometric Problems (25-83). Emerald Group Publishing Limited.

Strobl, R.O. and Robillard, P.D., 2008. Network design for water quality monitoring of surface freshwaters: A review. Journal of Environmental Management, 87(4), 639-648.

Sun, W., Xia, C., Xu, M., Guo, J. and 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.

Sundaray, S.K., Panda, U.C., Nayak, B.B. and Bhatta, D., 2006. Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of the Mahanadi river–estuarine system (India)–a case study. Environmental Geochemistry and Health, 28(4), 317-330.

Sutadian, A.D., Muttil, N., Yilmaz, A.G. and 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.

Sutadian, A.D., Muttil, N., Yilmaz, A.G. and Perera, B.J.C., 2018. Development of a water quality index for rivers in West Java Province, Indonesia. Ecological Indicators, 85, 966-982.

Swamee, P.K. and Tyagi, A., 2000. Describing water quality with aggregate index. Journal of Environmental Engineering, 126(5), 451-455.

Talling, J.F., 2010. Potassium—a non-limiting nutrient in fresh waters?. Freshwater Reviews, 3(2), 97-104.

Taoufik, G., Khouni, I. and Ghrabi, A., 2017. Assessment of physico-chemical and microbiological surface water quality using multivariate statistical techniques: a case study of the Wadi El-Bey River, Tunisia. Arabian Journal of Geosciences, 10(7), 181.

Tapiero, H., Townsend, D.M. and Tew, K.D., 2003. Trace elements in human physiology and pathology. Copper. Biomedicine & pharmacotherapy, 57(9), 386-398.

Thakur, T., Rishi, M.S., Naik, P.K. and Sharma, P., 2016. Elucidating hydrochemical properties of groundwater for drinking and agriculture in parts of Punjab, India. Environmental Earth Sciences, 75(6), 467.

Thapa, R., Gupta, S., Reddy, D.V. and Kaur, H., 2017. An evaluation of irrigation water suitability in the Dwarka river basin through the use of GIS-based modelling. Environmental Earth Sciences, 76(14), 471

Thi Minh Hanh, P., Sthiannopkao, S., The Ba, D. and Kim, K.W., 2010. Development of water quality indexes to identify pollutants in Vietnam's surface water. Journal of Environmental Engineering, 137(4), 273-283.

Tiwari, A.K., De Maio, M., Singh, P.K. and Mahato, M.K., 2015. Evaluation of surface water quality by using GIS and a heavy metal pollution index (HPI) model in a coal mining area, India. Bulletin of Environmental Contamination and Toxicology, 95(3), 304-310.

Tiwari, A.K., Singh, A.K., Singh, A.K. and Singh, M.P., 2017. Hydrogeochemical analysis and evaluation of surface water quality of Pratapgarh district, Uttar Pradesh, India. Applied Water Science, 7(4), 1609-1623.

Tripathi, M. and 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.

Tsai, W.P., Chang, F.J. and Herricks, E.E., 2016. Exploring the ecological response of fish to flow regime by soft computing techniques. Ecological Engineering, 87, 9-19.

Tsakiris, V., Alexakis, D. and Tsihrintzis, V.A., 2017. Assessing the quality of bottled water brands using a new water quality index.

Tyagi, S., Sharma, B., Singh, P. and Dobhal, R., 2013. Water quality assessment in terms of water quality index. American Journal of Water Resources, 1(3), 34-38.

Varol, M. and Şen, B., 2009. Assessment of surface water quality using multivariate statistical techniques: a case study of Behrimaz Stream, Turkey. Environmental Monitoring and Assessment, 159(1-4), 543.

Varol, M., Gökot, B., Bekleyen, A. and Şen, B., 2012. Water quality assessment and apportionment of pollution sources of Tigris River (Turkey) using multivariate statistical techniques—a case study. River Research and Applications, 28(9), 1428-1438.

Vasanthavigar, M., Srinivasamoorthy, K., Vijayaragavan, K., Ganthi, R.R., Chidambaram, S., Anandhan, P., Manivannan, R. and Vasudevan, S., 2010. Application of water quality index for groundwater quality assessment: Thirumanimuttar sub-basin, Tamilnadu, India. Environmental Monitoring and Assessment, 171(1-4), 595-609.

Vega, M., Pardo, R., Barrado, E. and Debán, L., 1998. Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. Water Research, 32(12), 3581-3592.

Viet, N.T. and Bhargava, D.S., 1989. Water quality and management of the Saigon river in Ho Chi Minh city (Vietnam). Indian Journal of Environmental Health, 31(4), 321-0.

Wang, W., Ndungu, A.W., Li, Z. and Wang, J., 2017. Microplastics pollution in inland freshwaters of China: a case study in urban surface waters of Wuhan, China. Science of the Total Environment, 575, 1369-1374.

Wang, Y., Wang, P., Bai, Y., Tian, Z., Li, J., Shao, X., Mustavich, L.F. and Li, B.L., 2013. Assessment of surface water quality via multivariate statistical techniques: a case study of the Songhua River Harbin region, China. Journal of Hydro-Environment Research, 7(1), 30-40.

Ward, R.C., Loftis, J.C. and McBride, G.B., 1986. The "data-rich but information-poor" syndrome in water quality monitoring. Environmental Management, 10(3), 291-297.

Whittaker, G., Barnhart, B., Färe, R. and Grosskopf, S., 2015. Application of index number theory to the construction of a water quality index: aggregated nutrient loadings related to the areal extent of hypoxia in the northern Gulf of Mexico. Ecological Indicators, 49, 162-168.

Winter, T.C., 1999. Relation of streams, lakes, and wetlands to groundwater flow systems. Hydrogeology Journal, 7(1), 28-45.

Wu, J., Li, P., Qian, H. and Chen, J., 2015. On the sensitivity of entropy weight to sample statistics in assessing water quality: statistical analysis based on large stochastic samples. Environmental Earth Sciences, 74(3), 2185-2195.

Wu, J., Xue, C., Tian, R. and Wang, S., 2017. Lake water quality assessment: a case study of Shahu Lake in the semiarid loess area of northwest China. Environmental Earth Sciences, 76(5), 232.

Yan, F., Liu, L., Li, Y., Zhang, Y., Chen, M. and Xing, X., 2015. A dynamic water quality index model based on functional data analysis. Ecological Indicators, 57, 249-258.

Yaseen, Z.M., Ramal, M.M., Diop, L., Jaafar, O., Demir, V. and Kisi, O., 2018. Hybrid Adaptive Neuro-Fuzzy Models for Water Quality Index Estimation. Water Resources Management, 32(7), 2227-2245.

Yidana, S.M., Ophori, D. and Banoeng-Yakubo, B., 2008. A multivariate statistical analysis of surface water chemistry data—The Ankobra Basin, Ghana. Journal of Environmental Management, 86(1), 80-87.

Yim, O. and Ramdeen, K.T., 2015. Hierarchical cluster analysis: comparison of three linkage measures and application to psychological data. Quant. Methods. Psychol, 11, 8-21.

Yuan, X., Huang, H., Zeng, G., Li, H., Wang, J., Zhou, C., Zhu, H., Pei, X., Liu, Z. and Liu, Z., 2011. Total concentrations and chemical speciation of heavy metals in liquefaction residues of sewage sludge. Bioresource Technology, 102(5), 4104-4110.

Zahedi, S., 2017. Modification of expected conflicts between drinking water quality index and irrigation water quality index in water quality ranking of shared extraction wells using multi criteria decision making techniques. Ecological Indicators, 83, 368-379.

Zahedi, S., Azarnivand, A. and Chitsaz, N., 2017. Groundwater quality classification derivation using multi-criteria-decision-making techniques. Ecological Indicators, 78, 243-252.

Zandbergen, P.A. and Hall, K.J., 1998. Analysis of the British Columbia water quality index for watershed managers: a case study of two small watersheds. Water Qual. Res. J. Canada, 33(4).

Zeng, Y., Lai, Z., Yang, W. and Gu, B., 2017. Multivariate statistical analysis of heavy metals in water from the Pearl River watershed, South China. Fresen. Environ. Bull, 26(5), 3350-3357.

Zhang, C., 2007. Fundamentals of environmental sampling and analysis. John Wiley & Sons.

Zhang, Q., Li, Z., Zeng, G., Li, J., Fang, Y., Yuan, Q., Wang, Y. and Ye, F., 2009. Assessment of surface water quality using multivariate statistical techniques in red soil hilly region: a case study of Xiangjiang watershed, China. Environmental Monitoring and Assessment, 152(1), 123-131.

Zhang, Q., Zheng, Y., Singh, V.P., Xiao, M. and Liu, L., 2016. Entropy-based spatiotemporal patterns of precipitation regimes in the Huai River basin, China. International Journal of Climatology, 36(5), 2335-2344.

Zhang, Y., Li, Y., Li, Y., Da, W., Yu, M. and Quan, Q., 2018. Application of multivariate statistical methods in the assessment of water quality in selected locations in Jialing River basin in Guangyuan, China. Water Science and Technology: Water Supply.

Zhang, Y., Xu, M., Li, X., Qi, J., Zhang, Q., Guo, J., Yu, L. and Zhao, R., 2018. Hydrochemical Characteristics and Multivariate Statistical Analysis of Natural Water System: A Case Study in Kangding County, Southwestern China. Water, 10(1), 80.

Zhao, C., Yao, S., Ding, Y. and Wang, J., 2015. Detection of precipitation variability based on entropy over nearly 50 years in Xinjiang, northwestern China. Theoretical and Applied Climatology, 122(3-4), 609-618.

PUBLICATIONS

International Journal (Published)

1. Singh, K.R., Dutta, R., Kalamdhad, A.S. and Kumar, B., 2019. An investigation on water quality variability and identification of ideal monitoring locations by using entropy based disorder indices. Science of The Total Environment, 647, pp.1444-1455.
2. Singh, K.R., Dutta, R., Kalamdhad, A.S. and Kumar, B., 2019. Information entropy as a tool in surface water quality assessment. Environmental Earth Sciences, 78(1), p.15. (Page No. 1-12)
3. Singh, K.R., Dutta, R., Kalamdhad, A.S. and Kumar, B., 2018. Risk characterization and surface water quality assessment of Manas River, Assam (India) with an emphasis on the TOPSIS method of multi-objective decision making. Environmental Earth Sciences, 77(23), p.780. (Page No. 1-10)
4. Singh, K.R., Goswami, A.P., Kalamdhad, A.S. and Kumar, B., 2019. Development of irrigation water quality index incorporating information entropy. Environment, Development and Sustainability, pp.1-14.
5. Singh, K.R., Goswami, A.P., Kalamdhad, A.S. and Kumar, B., 2019. Assessment of surface water quality of Pagladia, Beki and Kolong rivers (Assam, India) using multivariate statistical techniques. International Journal of River Basin Management, pp.1-10.

International Journal (Under Review)

1. Singh, K.R., Dutta, R., Kalamdhad, A.S. and Kumar, B., 2018. A critical review of existing heavy metal contamination indices and proposing an approach for indexing by using information entropy. *Ecological Indicators* (Under review)
2. Singh, K.R., Goswami, A., Kalamdhad, A.S. and Kumar, B., 2018. Surface water quality and health risk assessment of Kameng river (Assam, India), *Applied Water Science.* (Under Review)
3. Singh, K.R., Goswami, A., Kalamdhad, A.S. and Kumar, B., 2018. Application of entropy weighted water quality index (EWQI) for assessment of surface water quality of Baralia and Puthimari river, Assam (India). *Environmental Monitoring and Assessment* (Under Review)