Groundwater Quality Assessment using Entropy and CRITIC Weights on Water Quality Index Methods and Multiple Linear Regression in South 24 Parganas, India

A thesis submitted towards partial fulfillment of the requirements for the degree of

MASTER OF ENGINEERING

in **Water Resources and Hydraulic Engineering**

Course affiliated to Faculty Council of Engineering and Technology

Jadavpur University

submitted by

RIMPA SOM

Exam Roll No: M6WRP24017

Under the guidance of

DR. SUBHASISH DAS

Associate Professor

School of Water Resources Engineering, Jadavpur University

School of Water Resources Engineering

M.E. (Water Resources and Hydraulic Engineering) course affiliated to

Faculty Council of Engineering and Technology

Jadavpur University

Kolkata 700032, India

2024

Declaration of Originality and Compliance of Academic Ethics

I hereby declare that this thesis contains a literature survey and original research workby the undersigned candidate, as part of my **Master of Engineering** in **Water Resources and Hydraulic Engineering (6 Semester)** under the Faculty Council of Interdisciplinary Studies, Law and management in Jadavpur University during academic session 2023-24.

All information in this document has been obtained and presented in accordance with academic rules and ethical conduct.

I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name	**:**	**Rimpa Som**
Exam Roll Number	**:**	**M6WRP24017**
Thesis Title	**:**	**Groundwater Quality Assessment using Entropy and CRITIC Weights on Water Quality Index Methods and Multiple Linear Regression in South 24 Parganas, India**
Signature with Date	**:**	

M.E. (Water Resources and Hydraulic Engg.) course affiliated to

Faculty Council of Engineering and Technology,

Jadavpur University

Kolkata, India

Certificate of Recommendation

This is to certify that the thesis entitled "**Groundwater Quality Assessment using Entropy and CRITIC Weights on Water Quality Index Methods and Multiple Linear Regression in South 24 Parganas, India"** is a bonafide work carried out by **Mrs. Rimpa Som** under my supervision and guidance for partial fulfillment of the requirement for the Post Graduate Degree of Master of Engineering in Water Resources and Hydraulic Engineering (6 Semester) during the academic session 2023-2024.

THESIS ADVISOR
Dr. Subhasish Das
Associate Professor
School of Water Resources Engineering
Jadavpur University

DIRECTOR
Prof. (Dr.) Asis Mazumdar
School of Water Resources Engineering
Jadavpur University

DEAN
Prof. (Dr.) Ratan Mondal
Faculty of Interdisciplinary Studies, Law & Management,
Jadavpur University

M.E. (Water Resources and Hydraulic Engg.) course affiliated to

Faculty Council of Engineering and Technology

Jadavpur University

Kolkata, India

CERTIFICATE OF APPROVAL **

This foregoing thesis is hereby approved as a credible study of an engineering subject carried out and presented in a manner satisfactorily to warrant its acceptance as a prerequisite to the degree for which it has been submitted. It is understood that by this approval the undersigned do not endorse or approve any statement made or opinion expressed or conclusion drawn therein but approve the thesis only for the purpose forwhich it has been submitted.

_____

Committee of Final Examination

for the evaluation of the thesis

_____

_____

** Only in case the thesis is approved.

ACKNOWLEDGEMENTS

I express my sincere gratitude to my supervisor; ***Dr. Subhasish Das,*** *Associate Professor of School of Water Resources Engineering, Jadavpur University under whose supervision and guidance this work has been carried out. It would have been impossible to carry out this thesis work with confidence without his wholehearted involvement, advice, support and constant encouragement throughout. He has not only helped me in my thesis work but also has given valuable advice to proceed further in my life.*

My heartfelt gratitude goes to ***Prof. (Dr.) Asis Mazumdar****, Director,* ***Prof. (Dr.) Pankaj Kumar Roy****,* ***Dr. Rajib Das****, and* ***Dr. Gourab Banerjee****, Faculty members of School of Water Resources Engineering, Jadavpur University for their support.*

I am highly indebted to ***Mr. Chinmoy Ranjan Das, PhD Scholar, Jadavpur University*** *for his unconditional support and constant encouragement during my work.*

I am also grateful to all the staff of the School of Water Resources Engineering of Jadavpur University for their help and support.

I am thankful to all my friends for their support.

My special acknowledgements are due to my beloved family, my husband, Dr. Ashes Datta, who has always been with me when I needed help more than ever, I would like to thank my parents, Mr. Rathindra Nath Som, and Mrs. Kalpana Som, my elder sister Mrs. Barnali Som Kundu, my nephew Ayush Kundu, my brother-in-law Mr. Arnab Kundu who have been my strength and inspiration during the study. I am also grateful to my mother-in-law Mrs. Mousumi Datta, my father-in-law Mr. Asim Kr. Datta and my sister-in-law Ms. Rima Datta for their constant support and encouragements.

Date : 31.05.2024

RIMPA SOM

Examination Roll No- M6WRP24017

ABSTRACT

An internationally recognized standard to determine the water quality of any groundwater resource is the water quality index or WQI. There is a decrease in the water levels of several coastal zones' existing groundwater sources. As a result, residents in coastal areas like South 24 Parganas, India, significantly rely on groundwater, which is gradually degrading. Therefore, it is imperative to identify the sources of groundwater contamination and oversee groundwater quality conditions. Consequently, the present research aims to address in detail the causes and influencing factors of groundwater contamination, as well as the status and monitoring of water quality. In order to do so, it makes use of incredibly helpful cutting-edge technologies like multiple linear regression (MLR) and groundwater quality index (GWQI), which help to simplify complex tasks in place of more labor-intensive methods. Besides that, it also uses MLR to develop a new WQI equation for the South 24 Parganas district and assess its suitability. To assess whether water in coastal areas was suitable for drinking, many weight-based water quality index (WQI) techniques were used. To ascertain WQI related to groundwater in particular coastal regions of eastern India, pre-existing weights such as entropy and criteria-importance through inter-criteria correlation (CRITIC) weights were employed. Though seawater intrusion has increasingly impacted groundwater quality, there have been few investigations on WQI in this case. Therefore, the goal of this study is to determine the efficacy of the WQI weighting strategy and to provide an effective WQI approach by comparing the current weights for the assessment of water quality. Here, 100 locations from which the initial WQI results were derived were used to assess seven groundwater quality parameters: pH, TDS (total dissolved solids), Tur (turbidity), Mn (manganese), Fe (iron), TH (total hardness), and Cl (chloride). Based on these findings, MLR was used to create a new WQI equation that only included the following six variables: TH (total hardness), Cl (chloride), Mn (manganese), Fe (iron), and Tur (turbidity). Among the samples, it indicates that 41% and 5% are outstanding to good for drinking. Model 3 (dataset ratio 80:20, R^2 = 0.994, RMSE = 2.8007) has the best accuracy in WQI prediction out of the 12 best fitted models. It was feasible to show using the paired t-test that the new WQI's results did not differ substantially from the original WQI's results. The study found that the significant values greater than 0.05 were obtained for both the absolute WQI values and the quality ranges, indicating the suitability of the new equation. One of the important benefits of lowering the number of variables for water quality monitoring is that it lowers the cost of monitoring, even when it comes to minimizing the eclipse effect. In conclusion, it is determined that developing an equation with fewer variables is crucial for keeping an eye on the water quality in the South 24 Parganas district.

CONTENT

Page No.

Declaration *i*

Certificate of Recommendation *ii*

Certificate of Approval *iii*

Acknowledgements *iv*

Abstract *v*

Content *vi*

List of Tables *viii*

List of Figures *ix*

Abbreviations *x*

1. CHAPTER I:

1.1 Introduction 1

1.1.1 General background 1

1.1.2 Groundwater pollution and its effects 2

1.1.3 Background of groundwater condition of the study area 3

1.1.4 Problem statement 3

1.2 Groundwater quality assessment technique 4

1.3 Significance of the study 5

1.4 Objective of the study 5

1.5 Structure of thesis 5

2. CHAPTER II:

2.1 Literature Review 6

3. CHAPTER III:

3.1 Study area 11

4. CHAPTER IV:

4.1 Methodology 13

4.1.1 Collection of data 13

4.1.2 Statistical Analysis 13

4.1.3 Groundwater quality index (GWQI) estimation 13

4.1.3.1 Computation of Entropy weight based WQI 13
4.1.3.2 Computation of CRITIC weight based WQI 15
4.1.3.3 Comparison of CRITIC and Entropy methods through Eclipsing and Ambiguity 17
4.1.4 MLR Model Preparation 17
4.1.4.1 Normality Test for dataset input combination 17
4.1.4.2 Correlation Matrix 17
4.1.4.3 Multiple Linear Regression Analysis 18
4.1.4.4 t-Test 18
4.1.4.5 Evaluation of model's performance and acceptability 19

5. CHAPTER V:

5.1 Results and Discussion 24
5.1.1 Statistical Analysis of individual parameter 24
5.1.2 Formation of Entropy weight WQI and CRITIC weight WQI model 26
5.1.2.1 Computation of Entropy weight WQI (E_NWQI) 26
5.1.2.2 Computation of CRITIC weight WQI (C_WWQI) 34
5.1.2.3 Comparison of CRITIC and Entropy methods through Eclipsing and Ambiguity 37
5.1.3 MLR Model Preparation 38
5.1.3.1 Input Combination for Model preparation 38
5.1.3.2 Dataset Distribution for model preparation 39
5.1.3.3 Identification of MLR based new G_{WQI} models 40
5.1.3.4 Validation of models by t-test 49
5.1.3.5 Accuracy check of model 52
5.1.4 Discussion 53

6. CHAPTER VI:

6.1 Conclusions 54
6.2 Limitations 55
References 56

List of Tables

		Page No.
Table 4.1	Classification range related to E_NWQI, C_WWQI	16
Table 4.2	Values of different water quality monitoring parameters at 100 sampling stations	21
Table 5.1	Statistical analysis and permissible values as per BIS guidelines	25
Table 5.2	Value of Measured concentration (MCN_{ji})	26
Table 5.3	Values of minimum concentration [$(MCN_{ji})_{min}$] and maximum concentration [$(MCN_{ji})_{max}$]	29
Table 5.4	Values of Normalized Concentration (NCN_{ji})	29
Table 5.5	Value of Entropy based weights (E_NW_i) and Information Entropy (IEi) of individual parameter	32
Table 5.6	Value of E_NWQI	32
Table 5.7	Values of Correlation Coefficient	34
Table 5.8	Values of standard deviation, information amount and weight of individual parameter	35
Table 5.9	Values of C_WWQI	35
Table 5.10	Values of significance in K-S and S-W Test	38
Table 5.11	Correlation coefficient of Spearman Correlation Matrix	39
Table 5.12	Model Summary	40
Table 5.13	Model wise Coefficient values	46
Table 5.14	Model wise t-test result	50
Table 5.15	Model wise RMSE values	52

List of Figures

		Page No.
Fig. 3.1	Location of the Study Area	11
Fig. 4.1	Flow diagram of methodology for water quality prediction modeling	20
Fig. 5.1	Graphical representation of Entropy WQI against Sample point	34
Fig. 5.2	Graphical representation of CRITIC WQI against Sample point	37
Fig. 5.3	Graphical representation of model wise RMSE value	52

List of Abbreviation

WQI	Water Quality Index
GWQI	Groundwater Quality Index
MLR	Multiple Linear Regression
CRITIC	Criteria-Importance Through Inter-criteria Correlation
TDS	Total Dissolved Solids
Tur	Turbidity
Mn	Manganese
Fe	Iron
Cl	Chloride
As	Arsenic
F	Fluoride
TH	Total Hardness
WHO	World Health Organisation
BIS	Bureau of Indian Standards
WQP	Water Quality Parameter
CRZ	Coastal Regulation Zone
STDV	Standard Deviation
M	Mean
MAX	Maximum
MIN	Minimum
VAR	Variance
MCN_{ji}	Measured concentration
$(MCN_{ji})_{max}$	Maximum concentration

$(MCN_{ji})_{min}$	Minimum concentration
I_{Ei}	Information entropy
Pr_{ji}	Probability occurrences of normalized concentration
E_NW_i	Entropy based weights
Qr_i	Quality rating
SLV_i	Standard limiting value
E_NWQI	Entropy weight based WQI
C_{Rji}	Correlation coefficient
CWp	CRITIC weight of each parameter
C_WWQI	CRITIC weight based WQI
RMSE	Root mean square error
A_{WQIi}	Actual WQI
P_{WQIi}	Predicted WQI
K–S	Kolmogorov-Smirnov
S-W	Shapiro-Wilk
df	Degree of Freedom
WBPHED	West Bengal Public Health Engineering Department
NTU	Nephelometric Turbidity Unit
Mg/l	Milligrams per litre

CHAPTER – I

1.1. Introduction

1.1.1 General Background

Groundwater and surface water are the two types of water that can fulfill a community's demands for water supply. Any freshwater that enters marshes, stream systems, or lakes is considered surface water. However, groundwater is found in subterranean aquifers, which are found below ground. The majority of groundwater comes from precipitation and snowmelt that seeps through the surrounding soil into the bedrock. This water will settle between the rock layers' fissures and cavities as it moves downward. The majority of the earth's groundwater is found within half a mile or less of the surface. The water will pool and begin to flow upward as soon as it reaches an impervious layer of the earth. When groundwater pools, it can emerge on the surface as groundwater-fed lakes and springs.

The primary distinction between groundwater and surface water is the quality of the water in each. Surface water may include significant levels of toxins due to air fallout and runoff; hence, extensive treatment is required before the water is suitable for use as a community's water supply. Surface water frequently contains chemical contaminants that build up through runoff. Even though groundwater usually has fewer pollutants than surface water, there are still some that may be present. These pollutants are absorbed through soil percolation and seepage. However, at least some of the toxins can be naturally removed from the water by filtering it via the sediment layers that are located below the water table.

As a result, the soil and rocks that groundwater passes through filter out the majority of germs, making groundwater less vulnerable to bacterial contamination than surface water. Groundwater needs less treatment before it can be used as drinking water since it contains less impurity. The fact that groundwater is more easily obtainable in times of drought is another factor contributing to its frequent preference over surface water. The majority of surface water may evaporate during a drought, which could be problematic for any industry that uses surface water as its main source of water. Groundwater is less expensive and simpler to treat than surface water since it usually contains fewer impurities. Although surface waters are typically located in lakes and streams, groundwater is easier to access because it may be reached through wells anywhere that water is required.

As, one of the most important sources of drinking water in the world is groundwater, continuous monitoring and management of groundwater quality are essential to achieving sustainable development goals. Although groundwater is a vital natural resource used by people all over the world, numerous activities are causing its quality to decline. Health

issues are getting worse everywhere because people are drinking tainted water (Jang et al. 2020).

1.1.2 Groundwater Pollution and Its Effects

The population is currently dealing with a number of social, economic, and environmental issues, which are raising concerns about the environment. The demand for water is rising due to the new norms of modern living and the rate at which the population is growing. This and other elements contribute to a concerning rise in water resource pollution, together with improper solid waste and urban and industrial effluent disposal. Controlling, reducing, and monitoring the pollution of these natural resources is therefore very important (Valentini et al. 2021).

The development and prosperity of the area, aquatic life, human health, and development are all significantly impacted by the consumption of contaminated water. To remove contaminants and microorganisms dangerous to human health, water must be treated and disinfected. The World Health Organization, the European Union, the US Environmental Protection Agency, the Bureau of Indian Standards, and other authorities have established drinking water quality standards to preserve the ideal concentration of minerals necessary to maintain the body's mineral homeostasis (WHO, 2012; BIS, 2012). According to the guidelines established by the BIS and the WHO, the greatest concentration of an ion that is permitted is known as the permissible limit, and the desired limit is the minimal concentration that is necessary. Thus, in order to preserve the mineral balance, it is imperative that the concentration of these ions remain within the designated range. Even with sufficient treatment, water that lacks certain minerals should not be consumed because while most of the substances in water are safe to drink, some are necessary for bodily functions. Numerous investigative findings demonstrate that a mineral deficit has a direct or indirect impact on human health. In order to prevent negative health effects, the current study work addresses the creation of a water quality index that evaluates the mineral balance of water resources (Mukate et al. 2019).

Nearly 60% of people on Earth reside in coastal regions, and as a result of growing urbanization, more groundwater is being used to supply daily water needs. Saline water intrusion and other human factors are made possible by excessive groundwater withdrawal. In coastal places, these variables worsen the quality of the groundwater (Jesuraja et al. 2021). In many parts of the world, people rely on groundwater for drinking since there are not enough other water sources, and these groundwater supplies are in poor condition. As a result, accurate water quality prediction and assessment are crucial for human consumption and for taking appropriate action in the event that the water becomes unsafe. The primary water quality parameters (WQPs) are analyzed for assessment; however, labor and expense constraints make it impractical to continuously monitor all parameters in groundwater in developing nations (Das and Das, 2024).

Groundwater must be treated in a way that maintains the raw water's physical and chemical properties in accordance with accepted standards in order to yield drinking water of the highest caliber. The properties of the raw water have a significant impact on how well the treatment process performs. It is desirable to be able to forecast the raw-water quality over time in order to optimize the treatment procedures and provide high-quality potable water in an affordable way. This would provide early notification of variations in raw water quality that call for changing the parameters of the process.

1.1.3 Background of Groundwater Condition of The Study Area

The West Bengal coastal district includes the South 24 Parganas district. The West Bengal coastal track and the Lower Ganga plain are composed of sediments that have accumulated in lacustrine, marine, and fluvial environments. The study region's primary aquifers' sandstone alluvium is made up of recent Pleistocene sediments and tertiary deposits from significant coastal districts like South 24 Parganas. In the South 24 Parganas region of India, populations rely too heavily on groundwater sources for drinking and other purposes, despite the fact that the water's quality is steadily declining (Jaydhar et al. 2022).

1.1.4 Problem Statement

Since water class is frequently described in terms of numerous water quality metrics, a number of statistical techniques have recently been created to accurately examine and explain data as a result of the growth in parameters for groundwater quality study (Das et al., 2023). The primary problem with groundwater is that it can be extremely challenging to restore the groundwater class when it is contaminated (Chen et al. 2018). When more parameters are involved in the water quality assessment, it becomes very difficult to clarify the results using a conventional approach, such as the analysis of individual parameters, because each parameter has a different quality class. Determining the type and severity of water pollution accurately is a challenging and intricate task. The intricacy of looking at a lot of variables is the biggest issue with monitoring the quality of the water. Because of this, applying the WQI technique is thought to be a very efficient mechanism. Big data on water quality can be transformed by WQI into an index score that can be used to make informed decisions regarding the condition of the water (Das et al. 2023).

It is problematic for common people to understand individual parameter analysis and their standard range for drinking to classify the category of water. So it is necessary for common people one entire water quality evaluation by which they can easily understand the quality of water. The WQI method is a useful tool for evaluating the drinking-quality of groundwater. Water quality management and control depend heavily on WQI evaluation and water quality sense (Mahapatra et al. 2012). Several factors that genuinely impact the water quality of a specific source can be taken into consideration when creating a model with regard to WQI utilizing a statistical technique like MLR. Reducing the number of variables in the WQI offers cost-effectiveness that may be used for future WQI evaluations in a particular area (Valentini et al. 2021). When more than one parameter needs to be taken

into account at the same time, the MLR method should be employed to determine how they relate to each other. It is also crucial in determining which variable has the biggest effect on WQI. The MLR model is used to forecast water quality in order to facilitate monitoring (Wang et al. 2020).

1.2 Groundwater Quality Assessment Technique

In order to perform multiple linear regression (MLR) analysis for continuous monitoring, two variables are necessary; thus, the independent and dependent variables. In this study, water quality index (WQI) scores were regarded as the dependent variable and the measured parameter values of the physicochemical water quality variables were considered as the independent variables. The water quality index is a simple, but yet intelligible rating score that provides the composite influence of various water quality variables in a given water body. The index number is normally measured against a relative scale to explain the quality of water based on categories ranging from zero to one hundred (Banda and Kumarasamy, 2020).

In the interest of simplifying the interpretation of water quality index (WQI) values, mostly to accommodate non-technical individuals, an index categorization schema was established. The classification mechanism is based on an increasing scale index and the advantage of this system is that it is identical to a normal percentage hierarchy, therefore the public can easily relate to its function and interpretation. Both models applicable to this study yield WQI values between zero and hundred. Accordingly, the WQI scores are categorized using classes ranging from one to five; with “Class 1” representing water of the highest degree of purity with a possible maximum score of hundred and vice versa, “Class 5” denotes water quality of the poorest degree with index scores nearing or equal to zero. In order to close gaps identified in some of the existing classification scales, appropriate mathematical functions with logical linguistic descriptors which include but are not limited to, “greater than,” “less than,” and “equal to,” are used to appraise WQI scores and respectively assigned them to the corresponding category (Banda and Kumarasamy, 2020). WQI consists of a number of water quality parameters, which are transformed into a common scale. Such transformations are carried out since the monitored water quality data have different units. These values of the parameters transformed to a common scale are known as sub- indices (Sutadian et al. 2016).

In general, the information gained from the WQIs can be used for the following purposes (Sutadian et al. 2016):

a) To provide an overall status of water quality to the water authorities and the wider community.
b) To study the impacts of regulatory policies and environmental programs on environmental quality.
c) To compare the water quality of different sources and sites, without making a highly technical assessment of the water quality data.

d) To assist policy makers and the public in avoiding subjective assessments and subsequent biased opinions.

One of the main steps of WQI is the weightage approach like the subjective approach and objective approach. The subjective weightage approach is based on the expert opinion; on the other hand objective weightage approach is based on the dataset. Entropy and CRITIC weight methods are objective weight methods which are called multi criteria decision making methods.

1.3 Significance of the Study

The purpose of the study is to find the present scenario of the water quality in this study area and also find the groundwater quality through the prediction model (Multiple Linear Regression approach).

1.4 Objectives of the Study

To achieve this significance, the following objectives have been set:

- To find the groundwater quality through Entropy weight based and CRITIC weight based WQI approaches.
- To compare between Entropy and CRITIC weight methods.
- Normality test of the dataset and correlation analysis.
- Preparation of input combination and dataset distribution ratio.
- MLR Model preparation and its validation.

1.5 Structure of Thesis

To fulfill the proposed objectives, the present thesis is structured as follows. Following brief background information, Part 1 outlines the problem statement, groundwater quality assessment technique, significance, objectives of the study and organization of the thesis. Part 2 serves as a general review of various researches related to this study. Part 3 reports a brief description of the study area. Part 4 depicts the methodology used to achieve the designed goal. The results of the study have been discussed more concisely and critically in Part 5. Finally, Part 6 of the thesis includes the concluding remarks and limitations.

CHAPTER – II

2.1 Literature Review

Das et al. (2023) applied the Water Quality Index (WQI) and Multiple Linear Regression (MLR) approaches in their study to find the present status and monitor groundwater quality, as groundwater quality was contaminated regularly in their study area. They mentioned that WQI and MLR approaches help to simplify complex work instead of conventional methods like individual parameter analysis. With the help of these approaches, they obtained efficient outcomes and suggested that a model with less number of parameters provides financial benefits. The findings related to this study may be useful for designers and officials to monitor and protect coastal groundwater resources in selected areas.

Das (2018) studied groundwater quality assessment for its suitability for agricultural and domestic use and statistical investigation of groundwater quality parameters using correlation coefficient and regression analysis method in their chosen study area of South 24 Parganas district, WB. Groundwater samples (pre-monsoon and post-monsoon periods) were collected from tube wells that collected water from 180 meters to 330 meters across the study area. The quality analysis is done by evaluating physical quality, chemical quality, graphical presentation of geochemical data and quality criteria for groundwater salinization and groundwater use. The chemical characteristics of the water samples showed small seasonal variations. They performed statistical analysis through correlation and regression analysis to monitor the water quality parameters and predicted response variables and evaluated the relationship between the physicochemical properties of groundwater specimens. The study concluded that specimens from both periods are suitable for domestic and irrigation use.

Govindaraju et al. (2022) applied the WQI approach in their study to identify the groundwater quality parameters for drinking purposes. Individual groundwater parameters in the study area were analyzed, and compared to BIS Guidelines. The extracted components indicated that the sources that exceeded the permissible limit were due to geology, agri-environment, precipitation, domestic sewage, and industrial waste. In the study, several groundwater samples were found to have very low water quality which is a major threat to the habitat. Based on the WQI index, they got some efficient outcomes of groundwater samples in pre and post-monsoon which were suitable for drinking purposes.

Adimalla (2021) in their study used an entropy weighted water quality index (EWQI) and a pollution indicator of groundwater to evaluate the quality of groundwater in a semi-arid region of India. Data on physicochemical characteristics were compared to drinking water quality criteria using the EWQI and PIG methodologies. After this study, they concluded that according to EWQI data, only 60% of groundwater samples were moderately suitable

for drinking, while the remaining 40% of samples were unfit for consumption and would need to be treated.

Valentini et al. (2021) studied water quality analysis at Mirim Lagoon and developed a new WQI equation through the MLR method. Comparing the T-test obtained the value of 0.0467 and mentioned that with the use of less no of water quality parameter, the cost of the analysis will be less. They concluded that with significant accuracy, one can forecast and track the Mirim Lagoon's water quality using the new WQI equation that this study modeled. Furthermore, it should be mentioned that other water resources with historical monitoring records can also benefit greatly from the broad application of this strategy.

Aouiti et al. (2020) studied groundwater quality analysis at the Hajeb Layoun-Jelma basin of Central Tunisia and analyzed groundwater quality suitability for domestic and agricultural purposes using geochemical methods and geographic information systems. Major parameters and the chemical data are compared with the WHO standards and national standards, which indicate the groundwater is suitable for drinking except for some sampling points where salinity concentrations are high. For water quality assessment, here entropy method is used with other water quality indexing methods. After the analysis, they concluded that the majority of samples of the basin are not appropriate for irrigation uses and the representation of the EWQI method is optimistic.

Zhang et al. (2021) assessed the quality of groundwater in Jiaokou Irrigation District, Guanzhong Basin, through the integrated weight water quality index (IWQI) approach. Using combined statistical, geostatistical, and hydrogeochemical methods, the chemistry of groundwater is also investigated. The IWQI approach was developed using the CRITIC and entropy-weighted methods, and it performed well in describing and assessing groundwater quality. The study's conclusions and outcomes can help us understand the mechanisms influencing the chemistry of the groundwater in Jiaokou Irrigation District and can also be utilized as a benchmark for other comparable areas of the world.

Das and Das (2023) investigated weight-based water quality index (WQI) methods to assess whether water in coastal regions is suitable for consumption. To determine the WQI connected to groundwater in certain coastal areas of eastern India, a combination of methods including relative importance, entropy, CRITIC, integrated, and a newly added technique based on the removal impacts of criteria weights were applied to a dataset of seven water quality parameters of 70 locations. ANOVA was used to compare the scores of the current WQI schemes with a 95% confidence level. The result shows that there isn't much of a difference, as seen by the significance level of 92.5% (> 5%). According to the study's findings, new weighing techniques can be applied in place of the aforementioned weighting methods to establish WQI in this field.

Uddin et al. (2022) focused on developing an improved WQI model for assessing coastal water quality considering study areas such as Cork Harbour, Ireland. The model combines standard WQI components and incorporates modifications to enhance objectivity and data-

driven analysis. It uses the XGBoost machine learning method to rank and select water quality indicators. In order to compare the calculated indices with those derived from an MLR approach, R^2 and RMSE have been used as measures of aggregation effectiveness. The study compared eight sub-index aggregation functions, finding that the weighted quadratic mean and unweighted arithmetic mean functions were the most effective. Overall, the model aims to reduce ambiguity and errors in evaluating water quality, providing a more reliable assessment tool.

Jianhua et al. (2011) used the entropy weighted water quality index approach to evaluate the groundwater quality in Jingyuan County, Ningxia, China. Twelve hydrochemical parameters were chosen for the evaluation. The study came to the conclusion that the most important factors impacting the quality of groundwater are the amounts of iodine, TH, iron, and TDS. The rational assessment results align with the filed investigation's findings, indicating that Jingyuan County's groundwater is suitable for human consumption.

Chenini and Khemiri (2009) presented a methodology in their study for characterizing groundwater quality in watersheds using hydrochemical data. It uses multiple linear regression and structural equation modeling to explore the composition of phreatic aquifer groundwater samples and the origin of water mineralization in the Maknassy Basin, central Tunisia. The integrated results provide a method to explain groundwater chemistry origin.

Sajib et al. (2023) studied the Groundwater Quality Index (GWQI) model, developed in Bangladesh, which integrates ten water quality indicators, including six heavy metals, from 38 sampling sites. The model uses six machine learning techniques to enhance precision in water quality assessment. The study found that groundwater quality ranged from 'good' to 'fair', with most indicators meeting Bangladesh's and the World Health Organization's standards. Artificial neural networks outperformed other ML models in predicting GWQI scores, demonstrating high sensitivity and low uncertainty.

Das and Das (2024) developed an effective model for WQI prediction using entropy and CRITIC weight-based WQI (EWQI and CWQI) and multi-layer perceptron artificial neural network (MLP-ANN) technique. 1000 water sampling datasets from eastern India were used, with 65% and 67% being excellent to good for drinking. Out of three fitting model, the CRITIC-WQI-MLP-ANN-II model provides the best accuracy. The research's findings can be used to plan for the future supply of clean drinking water.

Ewaid et al. (2018) examined water quality data from ten stations on the Tigris River in Baghdad in 2016. The WQI was calculated from 11 parameters, and a water quality model (WQM) was developed, indicating unsuitable drinking water. The study found significant differences in WQI values between months and stations, with higher WQI values in winter and spring and lower values in summer and autumn. The WQM model, based on stepwise MLR analysis, improved prediction and can aid in Tigris River management.

Mahapatra et al. (2012) studied that the quality of groundwater is essential for corrective actions because of pollution from urbanization, industry, and population increase. This research uses Q-mode principal component analysis as an empirical method to categorize water samples according to ten quality indicators. The classification aids engineers and planners in implementing groundwater contamination mitigation strategies. The accuracy of the non-parametric approach relies on the choice of parameters, but it evaluates the WQI well and can be used for online estimates.

Chen et al. (2018) used multivariate statistical techniques to analyze the groundwater quality in the Longyan basin of Fujian Province, South China. The dataset included ten parameters from five locations. The results showed that the groundwater was classified into two groups, reflecting interannual variability. Factor analysis revealed three factors accounting for 85.5% and 100% of the total variance. Discriminant analysis revealed most parameters responsible for temporal variation. This study highlights the usefulness of multivariate statistical techniques for effective groundwater quality management.

Chandra et al. (2017) investigated water quality parameters in Vijayawada, Krishna district, Andhra Pradesh, using the Weighted Arithmetic WQI method. It analyzes water samples from 90 stations during pre and post monsoon seasons, selecting 19 prime locations with maximum physiochemical properties. The aim is to determine the suitability of water for drinking purposes based on the estimated WQI.

Fagbote et al. (2014) tested groundwater in farm settlements in Western Nigeria for pH, conductivity, turbidity, total dissolved solids, sulphate, phosphate, nitrogen nitrate, chloride, alkalinity, total hardness, calcium, magnesium, sodium, potassium, and total coliform in 2008 and 2009. Results showed some wells were out of the recommended range for drinking water. The WQI ranked the bore hole as 'excellent' in 2009, but 'good' in the rainy season. The study also found a moderate relationship between the WQI and total coliform.

Mukate et al. (2019) identified the need for a universally accepted WQI that considers both desirable and permissible limits in drinking water. The Integrated Water Quality Index (IWQI) has been developed, which is flexible, unbiased, and easy to calculate. The IWQI is classified into five categories based on the concentration of cations, anions, and other parameters in groundwater samples. The results show that 2% of samples are excellent for drinking, 39% good, 43% marginal, 8% poor, and 8% unsuitable. The IWQI can be adjusted to alert water managers of potential human health concerns.

Jaydhar et al. (2024) revealed that coastal groundwater quality is crucial for sustainable development in coastal regions due to water scarcity. Heavy metal contamination is a global health hazard, with 27%, 32%, and 10% of the area having high, high, and low human health hazard indexes. The western part of the region has high concentrations of heavy metals, affecting groundwater pollution. The study suggests that the proposed approach could be a useful tool for predicting groundwater vulnerability in other regions.

Gorgij et al. (2017) evaluated groundwater quality in the Azarshahr Plain, Iran, using entropy theory and spatial autocorrelation. Results showed that bicarbonate was associated with maximum spatial autocorrelation and minimum entropy, leading to the highest effectiveness rate. The entropy weighted WQI classified groundwater quality into five categories: excellent, good, moderate, poor, and extremely poor.

Kumar et al. (2020) analyzed water quality data in their study, conducted in Kozhikode City, Kerala, India, using multiple linear regression, structural equation modeling, and adaptive neuro-fuzzy inference system (ANFIS) modeling. To develop a single MLR model, the researchers integrated data from many places and times. After combining the three data sets, they created an SEM for the combined data set. When compared to the MLR model, the ANFIS model—which employed TDS as the dependent variable performed better on the external data set.

Egbueri et al. (2020) integrated multiple water quality and pollution indices to assess the quality of drinking water sources in Ojoto, Nigeria. Water samples were tested for various properties, and various indices revealed high metal pollution and environmental risks. The entropy-weighted WQI showed poor quality, while the groundwater pollution index confirmed highly polluted samples. Quantification of contamination and pollution index of groundwater models identified human activities as the main contributors to water pollution. The study recommends lifestyle changes and adequate water management policies.

Sutadian et al. (2016) used WQIs since the 1960s to evaluate river water quality. However, there is no universally accepted method for developing WQIs. This paper reviews 30 existing WQIs based on four steps: parameter selection, sub-indices generation, parameter weights generation, and index computation. Seven WQIs were identified as most important for wider use. Government support is crucial for wider use. Local water quality experts' opinions are recommended, especially in the first three steps. Uncertainty and sensitivity analysis are also recommended to reduce uncertainty.

Banda and Kumarasamy (2020) aimed to develop a surrogate WQI as an alternative tool for monitoring water quality in South Africa. The surrogate WQI uses multivariate statistical techniques to determine proxy variables, including chlorophyll-a, electrical conductivity, pondus Hydrogenium, and turbidity. The model is considered scientifically stable and has a predictive pattern identical to the ideal graph. The proposed surrogate WQI can be applied in remote monitoring systems and can help capture spatial and temporal changes in South African river catchments.

Chidiac et al. (2023) studied that the WQI is a widely used tool for describing water quality, based on physical, chemical, and biological factors. It involves parameter selection, data transformation, weighting, and sub-index value aggregation. This review study explores WQI's development, progress, benefits, drawbacks, and recent studies. To expand, WQIs should be linked to scientific breakthroughs and sophisticated statistical methods.

CHAPTER – III

3.1 Study Area

The number of Parganas (divisions) of Kolkata's zamindari, which Mir Jafar gave to the East India Company in 1757, is the source of the district's name, 24 Parganas. North 24 Parganas and South 24 Parganas are the two districts that were created out of this one. On March 1st, 1986, the South 24 Parganas district was officially recognized as a district (District Environment Plan, S 24 Pgs, 2023). The present headquarters of the district is at Alipore.

The South 24 Parganas district lies between latitudes 22˚30'45" and latitudes 20˚29' North, and longitudes 89˚4'56" and 88˚3'45" East and is surrounded by the Hooghly River in the west, the Bay of Bengal in the south, Calcutta City and North 24 Parganas in the north, and east side is demarcated by Bangladesh and the Bidya and Matla rivers. Subject to Survey of India Topo Sheet Numbers 79B/2, 79B/3, 79B/4, 79B/5, 79B/6, 79B/10, 79B/11, 79B/12, 79B/15, 79B/16, 79C/1, 79C/2, 79C/5, 79C/6, 79C/9, and 79C/10, the district is covered.

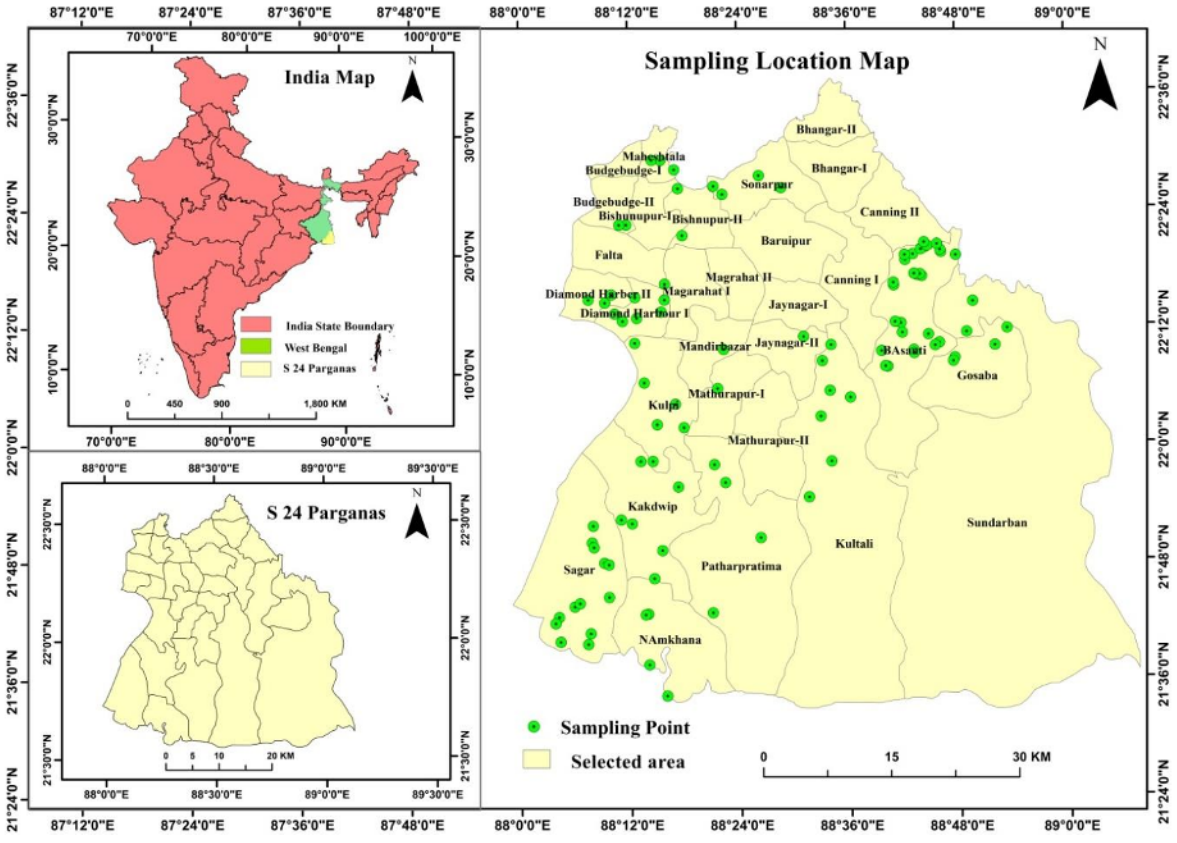


Fig. 3.1. Location of the study area highlighting sampling points

The district South 24-Parganas is distinguished by its closeness to both the pristine and breathtaking natural surroundings of the Sundarbans and the heavily developed metropolis of Kolkata on its eastern side. It holds a prominent place in West Bengal's industrial arena and is one of the district's most southern locations (District Survey Report, S 24 Pgs, 2018).

There are 29 Blocks, 31 Police Stations, 7 Municipalities and 5 Sub-Division in this district. The coast areas of the district stretch among the blocks Basanti, Canning – I, Canning- II, Diamond Harbour-I, Gosaba, Joynagar-II, Kakdwip, Kulpi, Kultali, Mathurapur-I, Mathurapur-II, Namkhana, Patharpratima, Sagar, which have long lines of coastal regulation zones (CRZs) within their boundaries.

In this study, more than 157 groundwater data have been collected from 129 Gram panchayats of 23 Blocks according to the availability of data. Out of those, 100 sampling points have been considered from 18 Blocks as per the dataset criteria of the study.

CHAPTER – IV

4.1 Methodology

4.1.1 Collection of Data

The dataset for this study includes more than 157 groundwater data samples from 124 Gram Panchayats of 13 Blocks of South 24 Parganas District were collected from the website of the Public Health Engineering Department, Government of West Bengal. The dataset contains WQP concentration of pH, TDS (Total Dissolved Solids), Turbidity, Mn (Manganese), Fe (Iron), Total hardness, Cl (Chloride), Arsenic (As) and Fluoride (F) in various blocks of South 24 Parganas sampled between 2020 to 2021. Fig. 3.1 shows the position of the sampling points. The magnitudes of the different groundwater quality monitoring parameters in the hundred sampling stations are detailed in Table 4.2. The approach in which this study is performed is sketched in Fig. 4.1.

4.1.2. Statistical Analysis

After the accumulation of the dataset, statistical analysis has been done. Basic descriptive statistics such as Mean (M), Variance (VAR), Standard Deviation (STDV), Maximum (MAX), and Minimum (MIN) of individual water quality parameters (WQP) have been calculated using Microsoft Excel to get an outline of the water quality. As per IS: 10500 (2012) guideline, standard values of individual parameters are compared with the value of the dataset (Drinking Water Specification, 2012) and checked to determine what number of samples fall above the standard limit.

4.1.3. Groundwater Quality Index (GWQI) Estimation

The Water Quality Index (WQI) is a very consistent, useful and effective method of evaluating and communicating information about the overall quality of water. (Gibrilla et al. 2011). The above-mentioned water quality parameters and standards are considered for each modeling method to find the WQI score for different location points. To evaluate the grade of water in the study area, two weight based water quality index modeling approaches i.e. Entropy weight based method (EW), and CRITIC weight based method (CW) are applied here.

4.1.3.1. Computation of Entropy Weight Based WQI

The EWQI is a representative measure of water quality that incorporates all physicochemical data into measurable values. As Arsenic (As) and Fluoride (F) are not

found in the accumulated dataset, so the other parameters are considered to calculate the Entropy weight based WQI in this research. A maximum number of parameters used in different studies (Das and Das, 2023; Das et al. 2023; Govindaraju et al. 2022; Gibrilla et al. 2011; Fagbote et al. 2014) related to this research. Recently, Das et al. (2023) suggested a WQI prediction model that utilizes most parameters as independent variables linked to the area of study. Therefore seven parameters have been incorporated that pertain to code availability, literature search and variable database availability into the WQI computation. The procedure to compute E_NWQI is as per the following steps described below (Adimalla, 2021).

The **first step** involves determining the entropy weight of a variable by taking into account the 's' number of water samples sampled (j = 1, 2, 3,....,s) and each sample consisting of 'P' water class variables (i = 1,2,3, ...,P). Initially, the eigenvalue matrix (E) is formed following Eq. (4.1).

$$\mathrm{E} = \begin{pmatrix} MCN_{11} & \cdots & MCN_{1P} \\ \vdots & \ddots & \vdots \\ MCN_{S1} & \cdots & MCN_{SP} \end{pmatrix} \tag{4.1}$$

The **second step** involves the normalized concentration ($\mathrm{NCN_{ji}}$) of each parameter being made using Eq. (4.2) and then the standard matrix (SM) is obtained using Eq. (4.3).

$$\mathrm{NCN_{ji}} = \frac{MCN_{ji} - (MCN_{ji})_{\min}}{(MCN_{ji})_{\max} - (MCN_{ji})_{\min}} \tag{4.2}$$

$$\mathrm{SM} = \begin{pmatrix} NCN_{11} & \cdots & NCN_{1P} \\ \vdots & \ddots & \vdots \\ NCN_{S1} & \cdots & NCN_{SP} \end{pmatrix} \tag{4.3}$$

where, $\mathrm{MCN_{ji}}$, $(\mathrm{MCN_{ji}})_{max}$, and $(\mathrm{MCN_{ji}})_{min}$ are the measured concentration, maximum concentration and minimum concentrations respectively of individual sample parameters.

The **third step** involves the information entropy ($\mathrm{I_{Ei}}$) equation which is obtained using Eq. (4.4).

$$\mathrm{I_{Ei}} = -\frac{1}{\ln(s)} \sum_{j=1}^{s} \mathrm{Pr}_{ji} \times \ln(\mathrm{Pr}_{ji}) \tag{4.4}$$

where, Pr_{ji} stands for Probability occurrences of normalized concentration of i^{th} parameter is evaluated by Eq. (4.5).

$$\mathrm{Pr}_{ji} = \frac{(1+NCN_{ji})}{\sum_{j=1}^{s}(1+NCN_{ji})} \tag{4.5}$$

Then entropy based weights (E_NW_i) are then calculated for the individual parameter by Eq. (4.6)

$$\mathrm{E_NW_i} = \frac{(1-I_{Ei})}{\sum_{j=1}^{s}(1-I_{Ei})} \tag{4.6}$$

The **fourth step** involves the quality rating (Qr_i) of the relevant parameters is calculated from the measured concentration (MCN_i) and standard limiting value (SLV_i) by Eq. (4.7)

$$\mathrm{Qr_i} = \frac{\frac{MCN_i}{SLV_i} \times 100}{\frac{MCN_{pH}-7}{SLV_{pH}-7}} \tag{4.7}$$

where, MCN_{pH} and SLV_{pH} correspond to measured concentrations and standard limiting values of pH as per BIS guidelines. If the MCN_{pH} value is greater than 7 and less than 7, the SLV_{pH} is considered to be 8.5 and 6.5 respectively, because the Qr_i value is positive**.**

In the **final step**, the Entropy weight based WQI value is evaluated by using Eq. (4.8).

$$\mathbf{E_NWQI} = \sum_{i=1}^{p} E_NW_i \times Qr_i \tag{4.8}$$

4.1.3.2. Computation of CRITIC Weight Based WQI

The full form of CRITIC is Criteria Importance Through Inter-criteria Correlation which is used in this study as an objective weighting method to calculate the relative weights of variables without relying on conventional information entropy flaws. This approach was proposed by Diakoulaki (Zhang et al., 2020; Zhang et al., 2021). The variables considered to calculate CWQI are the same as EWQI. The initial step in standardizing this database involves utilizing the CRITIC weighting method to compute the weight of each variable. The procedure to compute CWQI is as per the following steps described below by the equations (Das and Das, 2023).

The **first step** involves determining the Correlation coefficient (C_{Rji}) by the following Eq. (4.9).

$$C_{Rji} = \frac{\sum (MCN_{ji} - (MCN_{ji})_{AVG}) \times (NCN_{ji} - (NCN_{ji})_{AVG})}{\sqrt{\sum (MCN_{ji} - (MCN_{ji})_{AVG})^2 \times \sum (NCN_{ji} - (NCN_{ji})_{AVG})^2}} \quad (4.9)$$

where, $N1_{ji}$ and $N2_{ji}$ are normalized concentrations of two parameters, $(MCN_{ji})_{AVG}$ is the average of MCN_{ji} and $(NCN_{ji})_{AVG}$ is the average of $NCN_{ji.}$

The **second step** involves the determination of Information amount of the i^{th} parameter (I_{Fi}) using Eq. (4.10).

$$I_{Fi} = S_{DVi} \times \sum_{i=1}^{p} (1 - C_{Rji}) \quad (4.10)$$

where S_{DVi} stands for standard deviation related to the i^{th} parameter.

The **third step** is the calculation of the CRITIC weight of each parameter (CWp) using Eq. (4.11).

$$CWp = \frac{I_{Fi}}{\sum_{i=1}^{p} I_{Fi}} \quad (4.11)$$

The **fourth step** involves the computation of Quality rating (QRT_i) using Eq. (4.12).

$$QRT_i = \frac{(MCN_i - MCN_{Id})}{(SLV_i - MCN_{Id})} \times 100 \quad (4.12)$$

where MCN_i stands for Measured concentration, MCN_{Id} stands for Ideal concentration and SLV_i stands for suggested limiting value.

In the **final step**, CRITIC weight based WQI (C_WWQI) is evaluated by using Eq. (4.13).

$$\mathbf{C_W WQI} = \sum_{i=1}^{p} CW_p \times QRT_i \quad (4.13)$$

According to the WQI values from different approaches like Entropy and CRITIC weight, the quality of water has been classified as per the following table 4.1.

Table 4.1. Classification range related to E_NWQI, and C_WWQI (Das and Das, 2023).

Index score	Below 50	50–100	100–200	200–300	Above 300
Quality of water	Excellent to drink	Good for drinking	Poor	Very poor	Unsuited for drinking

4.1.3.3. Comparison of CRITIC and Entropy Method Through Eclipsing and Ambiguity

Concerns with eclipse and ambiguity are common with WQI techniques. If the water's final WQI score indicates that it is unacceptable, even though all parameters taken into consideration are within standard bounds, there may be ambiguity. Simultaneously, eclipsing might happen when certain WQPs are beyond their usual bounds, even though this isn't reflected in the final index score.

The above WQI approaches are associated with ambiguity and eclipsing since we see the related WQI scores after first identifying the sampling locations within the permitted range of all parameters. If the WQI score shows an undesirable status, then there is ambiguity; if not, there is not. Sajib et al. (2023) employed the method recommended by (Uddin et al. 2022) to examine the ambiguity and eclipsing error in their WQI model. Based on the range of parameters (0 to > 3) that deviate from the recommended values for eclipsing problem analysis, they employed four quality classes. In order to identify the eclipsing error, we also adhere to the same range of parameter numbers (0 to >3) exceeding the regulatory limits, but we create five quality classes because, in our study, WQI divides water into five classes. Excellent water quality (one parameter exceeds standard limit), good water quality (one parameter exceeds standard limit), bad water quality (two parameters exceed standard limit), extremely poor water quality (three parameters exceed standard limit), and unfit for drinking water (more than three parameters exceed standard limit).

4.1.4. MLR Model Preparation

4.1.4.1. Normality Test for Dataset Input Combination

Initially, to find the best input combination of the dataset, we first checked the dataset whether it was normal or not. First, an analysis was done on the sample distribution of the data that this study examined. Normality tests were run in order to determine this. The Kolmogorov-Smirnov (K–S) and Shapiro–Wilk (S–W) tests were performed at a significance level of 0.05 in order to confirm or deny the data's normality. In these tests, the distribution is assumed to be normal, and the null hypothesis (H_0) is rejected for p values less than 0.05. The K-S test is based on the greatest absolute difference in cumulative probabilities between the theoretical and empirical probabilities (Valentini et al. 2021). The S-W test's computation is done by the SPSS Tool.

4.1.4.2. Correlation Matrix

A correlation matrix with its coefficients and significance was employed following the analysis of the sample distribution to evaluate the correlation between the variables with the WQI and the chemical and microbiological variables under investigation. Because the

sample distribution of the data determines the correlation coefficient to be employed, normality checks were conducted beforehand. Spearman's coefficient, sometimes known as the Spearman Rhô coefficient, is a nonparametric coefficient that is advised for data lacking sample normality, while Pearson's coefficient is typically utilized for data that conforms to a normal distribution (Valentini et al. 2021).

4.1.4.3. Multiple Linear Regression Analysis

Using the multiple regression analysis (MLR) methodology, a novel WQI equation was modeled. Regression analysis enables a quantitative evaluation of the correlation between particular variables and an outcome of interest through this approach. All variables that show a significant correlation in the correlation matrix indicated above will be used to test regression models first, and the model equation with the best adjustments will then be identified. Using only the variables determined by the regression analysis, this equation will therefore be utilized to calculate a new water quality index for the water resource under study. The variables utilized in this work for the GWQI estimation and the MLR analysis are the same. Here, four ratios i.e. 70% and 30%, 75% and 25 %, 80 % and 20% and 85 % and 15 % of the total were selected at random for the purposes of model preparation and validation. This statistical method helps analyze how different independent parameters relate to a dependent parameter (Wu et al. 2020). That is stated in Eq. (4.14).

$$R = B_0 + B_1P_1 + B_2P_2 + \ldots\ldots\ldots + B_MP_M + R_E \quad (4.14)$$

where, P_1, P_M, and B_0 are the predictor parameters, R is the response parameter, and so on. Regression coefficients are denoted by B_M, and random error is denoted by R_E.

In order to create a suitable model, step-wise multiple regression analysis (MLR) was used to evaluate the relationship between the independent variables (groundwater quality parameters) and the dependent variable (GWQI) using IBM SPSS tool. The appropriate model was identified, and index scores were then computed. Finally, a paired samples t-test was used to identify the significant differences between the original GWQI and updated GWQI values.

4.1.4.4. t-Test

Student's t-test is used to determine whether there was a significant difference between the original WQI and the new WQI after obtaining the new WQI equation from the regression analysis and calculating the new index for each sample. This test will be performed on the quality ranges and numerical WQI results (Valentini et al. 2021). The significance of the difference between dependent sample pairs is determined using the t-test for paired samples. The use of the paired t-test is necessary because the results of the new WQI are dependent on the results of the old WQI, which were used to conduct the regression analysis. This test is a hypothesis test that seeks to assess a null hypothesis and an alternative, just like all

significance tests do. Though in this study, paired t-test is calculated using the SPSS tool, but in general paired *t*-test is calculated using Eq. (4.15).

$$t = \frac{\sum d}{\sqrt{\frac{n \times (\sum d^2) - (\sum d)^2}{n-1}}} \tag{4.15}$$

where, *d* stands for difference per paired value and n stands for number of samples.

The null hypothesis, which states that there is no significant variation between the old WQI and the new WQI, is rejected if, with 95% confidence, the p value is less than 0.05, according to this test. In other words, the original and new indices are not substantially different if p > 0.05 (Das et al., 2023; Valentini et al. 2021).

4.1.4.5. Evaluation of Models' Performance and Their Acceptability

The projected WQI values are compared with the original WQI dataset in order to assess the performance of the models. To determine the correctness of models, the following statistical formula for root mean square error (RMSE) is taken into consideration. If validated models are all accepted through t-test, then out of all models, the most accurate model is selected based on the lowest RMSE, as is customary in the literature. RMSE calculation is done using the following Eq. (4.16) (Elbeltagi et al. 2020).

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(A_{WQIi} - P_{WQIi})^2} \tag{4.16}$$

where, A_{WQIi} stands for actual WQI and P_{WQIi} stands for predicted WQI value.

FLOW DIAGRAM

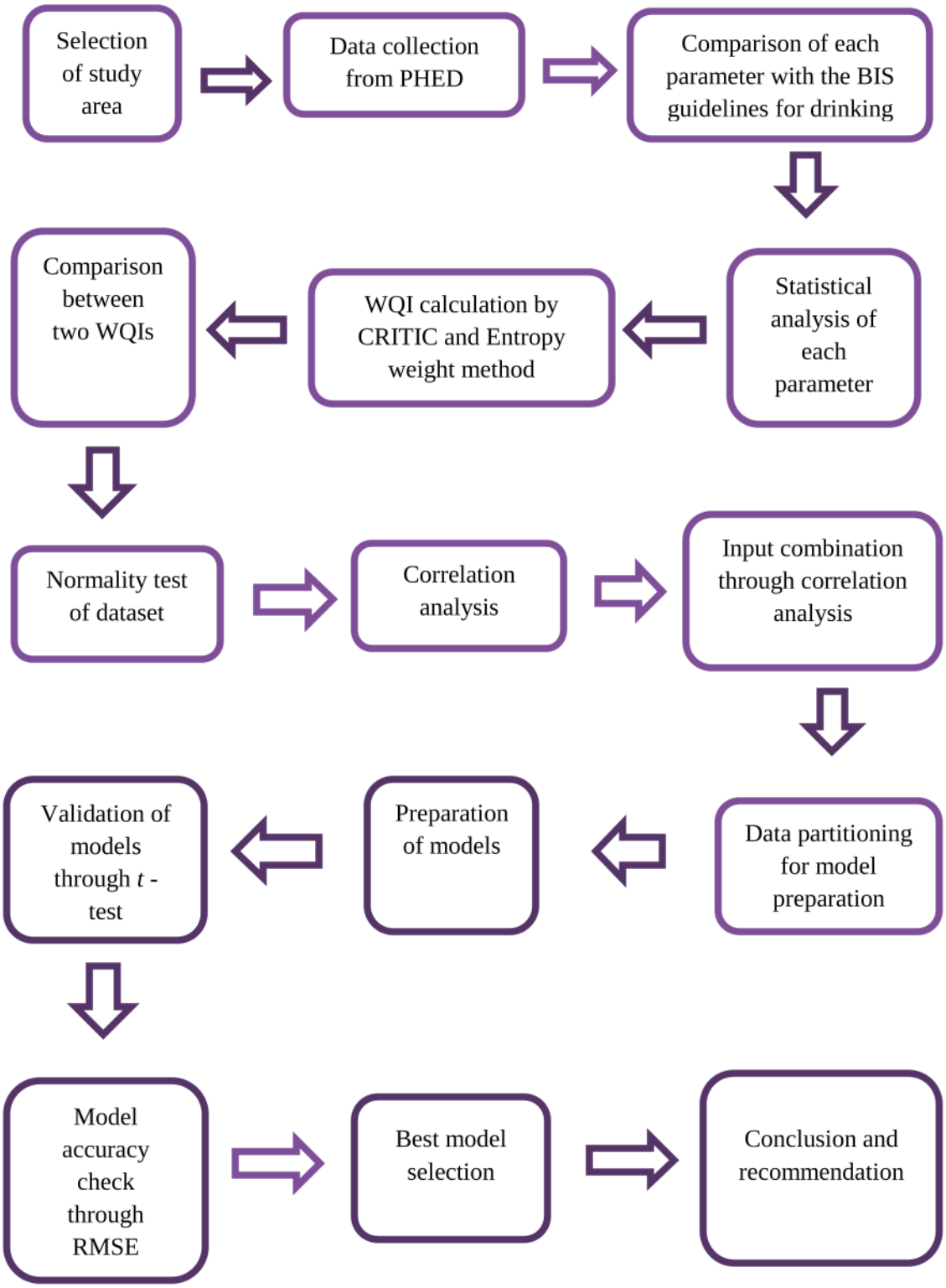


Fig. 4.1. Flow diagram of methodology for water quality prediction modelling.

Table 4.2. Values of different water quality monitoring parameters at 100 sampling stations.

Sample number	pH	TDS (mg/l)	Turbidity (NTU)	Iron (mg/l)	Manganese (mg/l)	Total Hardness (mg/l)	Chloride (mg/l)
1	7.44	792	0.65	0.26	0.25	124	192
2	7.26	1347	2.17	0.18	0.25	148	198
3	7.62	742	0.52	0.23	0.25	176	246
4	8.1	692	2.1	0.23	0.25	152	226
5	7.49	792	0.81	0.14	0.25	126	126
6	7.1	1645	4.25	0.38	0.25	218	345
7	7.25	897	0.83	0.18	0.25	182	326
8	7.65	1362	2.79	0.15	0.25	132	262
9	7.38	857	1.36	0.16	0.25	132	165
10	7.93	925	1.15	0.14	0.25	148	210
11	7.32	845	0.92	0.16	0.25	142	216
12	7.64	1410	4.32	0.32	0.25	184	292
13	7.81	1172	1.81	0.14	0.25	144	182
14	7.25	963	1.32	0.2	0.25	132	174
15	7.15	856	1.24	0.14	0.25	142	196
16	7.35	894	1.37	0.16	0.25	132	184
17	7.42	946	0.89	0.1	0.25	152	216
18	7.61	924	1.32	0.15	0.25	136	218
19	8.2	1327	2.14	0.24	0.25	174	248
20	7.1	856	1.32	0.15	0.25	156	208
21	7.83	795	0.71	0.18	0.18	136	187
22	7.64	684	0.71	0.14	0.25	158	192
23	7.85	1682	4.16	0.36	0.25	218	345
24	8.1	872	1.25	0.32	0.25	136	198
25	7.36	945	0.87	0.23	0.25	156	270
26	7.44	1064	1.47	0.22	0.25	145	210
27	7.32	1064	1.87	0.18	0.25	162	196
28	8.05	892	0.58	0.09	0.25	152	224
29	7.82	1349	3.42	0.12	0.25	175	264
30	7.56	1234	2.34	0.2	0.25	134	206
31	7.55	1659	3	0.28	0.05	200	200
32	7.76	496	1.56	0.22	0.17	320	70.9
33	8.03	661	0.79	0.16	0.09	200	142
34	7.59	829	0.7	0.18	0.01	230	270
35	7.34	775	1.48	0.25	0.36	190	213

Sample number	pH	TDS (mg/l)	Turbidity (NTU)	Iron (mg/l)	Manganese (mg/l)	Total Hardness (mg/l)	Chloride (mg/l)
36	7.39	711	0.12	0.21	0.02	230	234
37	7.31	732	0.01	0.14	0.16	250	234
38	7.39	999	0.53	0.27	0.21	300	341
39	7.25	1160	0.63	0.2	0	370	375
40	7.34	1376	0.68	0.28	0.34	390	533
41	7.36	984	0.08	0.19	0.01	250	433
42	7.12	1125	1.05	0.26	0.79	410	346
43	7.62	840	2.28	3.93	0.11	200	241.06
44	7.49	837	0.75	0.3	0.25	134	187
45	7.35	1068	1.35	0.23	0.25	175	238
46	7.46	794	0.73	0.12	0.25	149	258
47	8.1	1526	3.46	0.38	0.25	210	348
48	7.22	972	0.92	0.22	0.25	164	272
49	7.81	856	0.93	0.93	0.25	178	314
50	7.43	617	0.68	0.43	0.087	208	146
51	7.45	586	0.52	0.37	0.154	112	224
52	7.36	720	0.84	0.42	0.15	160	184.6
53	7.43	1115	0.66	0.63	0.16	180	298.2
54	7.53	570	0.81	0.6	0.15	164	184.6
55	7.62	822	0.48	0.36	0.13	144	198.8
56	8.08	703	0.55	0.439	0.028	116	178.92
57	7.71	571	0.43	0.438	0.056	164	159.04
58	7.9	561	0.4	0.586	0.017	104	201.64
59	8.06	1013	0.89	0.761	0.133	236	187.44
60	7.8	904	0.89	0.629	0.13	188	173.24
61	7.69	875	0.45	0.363	0.073	156	176.08
62	7.72	666	0.69	0.439	0.05	120	156.2
63	7.77	732	0.57	0.535	0.035	188	164.72
64	7.84	541	0.62	0.402	0.056	92	190.28
65	7.28	525	0.89	0.551	0.062	136	122.12
66	7.39	771	0.81	0.525	0.045	160	150.52
67	7.68	601	0.77	0.354	0.027	104	227.2
68	7.78	591	0.91	0.297	0.2519	272	148.89
69	7.81	1099	0.74	0.05	0.25	104	347.5
70	7.79	657	0.96	0.109	0.063	88	139.16
71	7.98	643	0.83	0.091	0.038	76	113.6
72	7.98	547	0.51	0.098	0.041	68	105.08

Sample number	pH	TDS (mg/l)	Turbidity (NTU)	Iron (mg/l)	Manganese (mg/l)	Total Hardness (mg/l)	Chloride (mg/l)
73	7.87	532	0.63	0.081	0.039	64	102.24
74	7.92	537	0.46	0.031	0.013	68	96.4
75	7.96	669	0.62	0.065	0.062	76	130.6
76	7.6	1066	0.74	0.23	0.02	96	490.06
77	7.54	639	1.67	0.19	0.02	48	217.8
78	7.68	731	0.26	0.13	0.02	64	272.26
79	7.73	857	1.32	0.2	0.03	44	299.48
80	7.41	385	0.46	0.95	0.04	119	87
81	7.63	390	0.29	0.06	0.05	124	65
82	7.77	428	0.42	0.19	0.06	156	136
83	7.86	399	0.8	0.07	0.13	118	85
84	7.9	403	0.35	0.09	0.03	84	63
85	7.81	372	0.17	0.05	0.17	80	69
86	7.87	408	0.02	0.02	0.04	84	59
87	7.67	368	0.43	0.27	0.04	78	56
88	7.87	407	0.29	0.04	0.01	100	95
89	7.53	378	0.29	0.08	0.07	144	99
90	7.39	426	0.7	0.04	0.04	144	99
91	7.56	453	2	0.36	0.07	120	102
92	7.59	422	1.4	0.09	0.02	84	75
93	7.53	1460	3.12	0.338	0	141	383.71
94	7.63	1060	2.17	1.203	0	151	387.46
95	7.52	1420	3.19	0.402	0	129	387.06
96	7.32	1120	6.59	1.175	0	169	314.19
97	7.65	518	4.92	0.44	0.05	224	108
98	7.55	420	2.87	0.12	0.27	168	39
99	7.37	606	1.21	0.33	0.05	240	210
100	7.48	600	0.61	0.24	0.16	180	118

CHAPTER – V

5.1 Results and Discussion

As it is discussed earlier the dataset containing all the drinking water parameters of various blocks of South 24 Parganas was collected from the WBPHED website through the link https://maps.wbphed.gov.in/webgis/?district=&block=&gp=&scheme=&search=Show. Out of all the datasets, seven parameters have been chosen from 100 sampling points. The parameters are pH, TDS, turbidity, iron, manganese, total hardness and chloride.

5.1.1 Statistical Analysis of Individual Parameters

To obtain a broad perspective and identify differences within the datasets, basic statistics were computed for each groundwater dataset (Das et al. 2023). After the preparation of the dataset, descriptive statistics of each parameter have been calculated by evaluating mean, variance, standard deviation, minimum value and maximum value. Variables related to groundwater quality are thought to be the most important foundation for highlighting the nature, diversity, and character of groundwater. Details of statistical indicators, and Standard value of every parameter and the number of samples above the standard are mentioned in Table 5.1.

In the statistical analysis, it is found that the variance of TDS and Chloride is high, out of them TDS is maximum. The minimum, maximum and mean values of TDS are 368, 1682 and 831.29 mg/l respectively. The permissible value of TDS is 500 mg/l. It is found that in the study area, out of 100 samples, 85 numbers of samples have a TDS value more than the permissible value.

The average concentration of pH in groundwater in the study area is 7.61, indicating a slightly acidic to alkaline type, with a pH range of 7.10 - 8.20. As per IS:10500 (2012), the permissible range of pH is 6.5 to 8.5. In the study areas, all the values of pH are within the permissible range.

The permissible value of turbidity is 1 NTU. In the study area, out of 100 samples, 38 nos samples have turbidity conc. above the permissible limit. The minimum, maximum and mean values of turbidity are found 0.01, 6.59 and 1.2561 NTU respectively.

Iron is a very important element that has an impact on human health; either too much or too little iron intake can lead to a number of disorders (Das et al. 2023). The Iron range is between 0.02 mg/l to 3.93 mg/l with an average concentration of 0.32 mg/l in the study area. 33% of the samples exceeded the permissible limit of Iron (Fe) i.e. 0.3 mg/l for drinking given by IS 10500 (Drinking Water Specification 2012).

Naturally occurring in groundwater, manganese is mostly found in anaerobic conditions. Its concentration is influenced by various factors such as the lithologic state of the aquifer, the flow paths of groundwater, and the composition of rainfall (Ram et al., 2021). The concentration of Manganese ranges between 0 to 0.79 mg/l with a mean of 0.15 mg/l. The permissible value of Manganese is 0.1 mg/l. In the study area, almost 56% of samples exceed the permissible limit.

Water with a TH value (mg/l) between 150 and 300 and beyond 300 is considered hard or very hard, respectively (Todd and Mays, 2005). The average concentration of Total hardness in groundwater in the study area is 157.91 mg/l, indicating a range of 44 – 410 mg/l. The permissible value of Total Hardness is 200 mg/l specified by IS:10500 (2012). In this study, 17 % of samples exceed the permissible limit.

The chloride range is between 39 mg/l to 533 mg/l with an average concentration of 207.96 mg/l in the study area. The permissible value of chloride is 250 mg/l. In this study, chloride concentrations of 26 samples are above the permissible limit.

Table 5.1. Statistical analysis and permissible values as per BIS guidelines.

WQP	Mean	Standard Deviation	Variance	Minimum	Maximum	Standard Value as per IS:10500 (2012)	Number of Samples Falling above Standard
pH	7.608	0.260	0.068	7.1	8.2	6.5 - 8.5	0
TDS (mg/l)	831.29	321.217	103180.2	368	1682	500	85
Turbidity (NTU)	1.256	1.146	1.313	0.01	6.59	1	38
Iron (mg/l)	0.316	0.428	0.183	0.02	3.93	0.3	33
Manganese (mg/l)	0.149	0.122	0.015	0	0.79	0.1	56
Total Hardness (mg/l)	157.91	66.621	4438.305	44	410	200	17
Chloride (mg/l)	207.961	97.324	9471.874	39	533	250	26

5.1.2. Formation of Entropy Weight WQI and CRITIC Weight WQI Model

In this study, two indexing methods have been used to determine the quality of groundwater associated with drinking i.e. Entropy weight WQI and CRITIC weight WQI.

5.1.2.1. Computation of Entropy Weight WQI (E_NWQI)

The WQI with regard to entropy weight is often used to measure the drinking water class. The greatest and lowest weights of the characteristics derived from the entropy technique have both been identified as influencing factors in water grades (Jianhua et al. 2011). E_NWQI of each sample are estimated based on the method described above. The E_NWQI value varies from 43.1257 to 213.9703. To estimate the E_NWQI value, all the required data are calculated step by step and given below in tabular form.

Table 5.2. Values of measured concentrations (MCN_{ji}).

MCNji							
Sample Number	**pH**	**TDS (mg/l)**	**Turbidity (NTU)**	**Iron (mg/l)**	**Manganese (mg/l)**	**Total Hardness (mg/l)**	**Chloride (mg/l)**
1	7.44	792	0.65	0.26	0.25	124	192
2	7.26	1347	2.17	0.18	0.25	148	198
3	7.62	742	0.52	0.23	0.25	176	246
4	8.1	692	2.1	0.23	0.25	152	226
5	7.49	792	0.81	0.14	0.25	126	126
6	7.1	1645	4.25	0.38	0.25	218	345
7	7.25	897	0.83	0.18	0.25	182	326
8	7.65	1362	2.79	0.15	0.25	132	262
9	7.38	857	1.36	0.16	0.25	132	165
10	7.93	925	1.15	0.14	0.25	148	210
11	7.32	845	0.92	0.16	0.25	142	216
12	7.64	1410	4.32	0.32	0.25	184	292
13	7.81	1172	1.81	0.14	0.25	144	182
14	7.25	963	1.32	0.2	0.25	132	174
15	7.15	856	1.24	0.14	0.25	142	196
16	7.35	894	1.37	0.16	0.25	132	184
17	7.42	946	0.89	0.1	0.25	152	216
18	7.61	924	1.32	0.15	0.25	136	218
19	8.2	1327	2.14	0.24	0.25	174	248
20	7.1	856	1.32	0.15	0.25	156	208
21	7.83	795	0.71	0.18	0.18	136	187
22	7.64	684	0.71	0.14	0.25	158	192

MCNji							
Sample Number	**pH**	**TDS (mg/l)**	**Turbidity (NTU)**	**Iron (mg/l)**	**Manganese (mg/l)**	**Total Hardness (mg/l)**	**Chloride (mg/l)**
23	7.85	1682	4.16	0.36	0.25	218	345
24	8.1	872	1.25	0.32	0.25	136	198
25	7.36	945	0.87	0.23	0.25	156	270
26	7.44	1064	1.47	0.22	0.25	145	210
27	7.32	1064	1.87	0.18	0.25	162	196
28	8.05	892	0.58	0.09	0.25	152	224
29	7.82	1349	3.42	0.12	0.25	175	264
30	7.56	1234	2.34	0.2	0.25	134	206
31	7.55	1659	3	0.28	0.05	200	200
32	7.76	496	1.56	0.22	0.17	320	70.9
33	8.03	661	0.79	0.16	0.09	200	142
34	7.59	829	0.7	0.18	0.01	230	270
35	7.34	775	1.48	0.25	0.36	190	213
36	7.39	711	0.12	0.21	0.02	230	234
37	7.31	732	0.01	0.14	0.16	250	234
38	7.39	999	0.53	0.27	0.21	300	341
39	7.25	1160	0.63	0.2	0	370	375
40	7.34	1376	0.68	0.28	0.34	390	533
41	7.36	984	0.08	0.19	0.01	250	433
42	7.12	1125	1.05	0.26	0.79	410	346
43	7.62	840	2.28	3.93	0.11	200	241.06
44	7.49	837	0.75	0.3	0.25	134	187
45	7.35	1068	1.35	0.23	0.25	175	238
46	7.46	794	0.73	0.12	0.25	149	258
47	8.1	1526	3.46	0.38	0.25	210	348
48	7.22	972	0.92	0.22	0.25	164	272
49	7.81	856	0.93	0.93	0.25	178	314
50	7.43	617	0.68	0.43	0.087	208	146
51	7.45	586	0.52	0.37	0.154	112	224
52	7.36	720	0.84	0.42	0.15	160	184.6
53	7.43	1115	0.66	0.63	0.16	180	298.2
54	7.53	570	0.81	0.6	0.15	164	184.6
55	7.62	822	0.48	0.36	0.13	144	198.8
56	8.08	703	0.55	0.439	0.028	116	178.92
57	7.71	571	0.43	0.438	0.056	164	159.04
58	7.9	561	0.4	0.586	0.017	104	201.64
59	8.06	1013	0.89	0.761	0.133	236	187.44
60	7.8	904	0.89	0.629	0.13	188	173.24
61	7.69	875	0.45	0.363	0.073	156	176.08

MCNji							
Sample Number	**pH**	**TDS (mg/l)**	**Turbidity (NTU)**	**Iron (mg/l)**	**Manganese (mg/l)**	**Total Hardness (mg/l)**	**Chloride (mg/l)**
62	7.72	666	0.69	0.439	0.05	120	156.2
63	7.77	732	0.57	0.535	0.035	188	164.72
64	7.84	541	0.62	0.402	0.056	92	190.28
65	7.28	525	0.89	0.551	0.062	136	122.12
66	7.39	771	0.81	0.525	0.045	160	150.52
67	7.68	601	0.77	0.354	0.027	104	227.2
68	7.78	591	0.91	0.297	0.2519	272	148.89
69	7.81	1099	0.74	0.05	0.25	104	347.5
70	7.79	657	0.96	0.109	0.063	88	139.16
71	7.98	643	0.83	0.091	0.038	76	113.6
72	7.98	547	0.51	0.098	0.041	68	105.08
73	7.87	532	0.63	0.081	0.039	64	102.24
74	7.92	537	0.46	0.031	0.013	68	96.4
75	7.96	669	0.62	0.065	0.062	76	130.6
76	7.6	1066	0.74	0.23	0.02	96	490.06
77	7.54	639	1.67	0.19	0.02	48	217.8
78	7.68	731	0.26	0.13	0.02	64	272.26
79	7.73	857	1.32	0.2	0.03	44	299.48
80	7.41	385	0.46	0.95	0.04	119	87
81	7.63	390	0.29	0.06	0.05	124	65
82	7.77	428	0.42	0.19	0.06	156	136
83	7.86	399	0.8	0.07	0.13	118	85
84	7.9	403	0.35	0.09	0.03	84	63
85	7.81	372	0.17	0.05	0.17	80	69
86	7.87	408	0.02	0.02	0.04	84	59
87	7.67	368	0.43	0.27	0.04	78	56
88	7.87	407	0.29	0.04	0.01	100	95
89	7.53	378	0.29	0.08	0.07	144	99
90	7.39	426	0.7	0.04	0.04	144	99
91	7.56	453	2	0.36	0.07	120	102
92	7.59	422	1.4	0.09	0.02	84	75
93	7.53	1460	3.12	0.338	0	141	383.71
94	7.63	1060	2.17	1.203	0	151	387.46
95	7.52	1420	3.19	0.402	0	129	387.06
96	7.32	1120	6.59	1.175	0	169	314.19
97	7.65	518	4.92	0.44	0.05	224	108
98	7.55	420	2.87	0.12	0.27	168	39
99	7.37	606	1.21	0.33	0.05	240	210
100	7.48	600	0.61	0.24	0.16	180	118

Table 5.3. Values of minimum concentrations [$(MCN_{ji})_{min}$] and maximum concentrations [$(MCN_{ji})_{max}$].

WQP	pH	TDS (mg/l)	Turbidity (NTU)	Iron (mg/l)	Manganese (mg/l)	Total Hardness (mg/l)	Chloride (mg/l)
$(MCN_{ji})_{min}$	7.1	368	0.01	0.02	0	44	39
$(MCN_{ji})_{max}$	8.2	1682	6.59	3.93	0.79	410	533

Normalized concentration is calculated from measured concentration by using the above mentioned formula (Eq. 4.2), which is given below in the tabular form.

Table 5.4. Values of Normalized Concentration (NCN_{ji}).

NCNji							
Sample No.	pH	TDS (mg/l)	Turbidity (NTU)	Fe (mg/l)	Mn (mg/l)	Total hardness (mg/l)	Cl (mg/l)
1	0.3091	0.3227	0.0973	0.0614	0.3165	0.2186	0.3097
2	0.1455	0.7451	0.3283	0.0409	0.3165	0.2842	0.3219
3	0.4727	0.2846	0.0775	0.0537	0.3165	0.3607	0.4190
4	0.9091	0.2466	0.3176	0.0537	0.3165	0.2951	0.3785
5	0.3545	0.3227	0.1216	0.0307	0.3165	0.2240	0.1761
6	0.0000	0.9718	0.6444	0.0921	0.3165	0.4754	0.6194
7	0.1364	0.4026	0.1246	0.0409	0.3165	0.3770	0.5810
8	0.5000	0.7565	0.4225	0.0332	0.3165	0.2404	0.4514
9	0.2545	0.3721	0.2052	0.0358	0.3165	0.2404	0.2551
10	0.7545	0.4239	0.1733	0.0307	0.3165	0.2842	0.3462
11	0.2000	0.3630	0.1383	0.0358	0.3165	0.2678	0.3583
12	0.4909	0.7930	0.6550	0.0767	0.3165	0.3825	0.5121
13	0.6455	0.6119	0.2736	0.0307	0.3165	0.2732	0.2895
14	0.1364	0.4528	0.1991	0.0460	0.3165	0.2404	0.2733
15	0.0455	0.3714	0.1869	0.0307	0.3165	0.2678	0.3178
16	0.2273	0.4003	0.2067	0.0358	0.3165	0.2404	0.2935
17	0.2909	0.4399	0.1337	0.0205	0.3165	0.2951	0.3583
18	0.4636	0.4231	0.1991	0.0332	0.3165	0.2514	0.3623
19	1.0000	0.7298	0.3237	0.0563	0.3165	0.3552	0.4231
20	0.0000	0.3714	0.1991	0.0332	0.3165	0.3060	0.3421
21	0.6636	0.3250	0.1064	0.0409	0.2278	0.2514	0.2996
22	0.4909	0.2405	0.1064	0.0307	0.3165	0.3115	0.3097
23	0.6818	1.0000	0.6307	0.0870	0.3165	0.4754	0.6194
24	0.9091	0.3836	0.1884	0.0767	0.3165	0.2514	0.3219
25	0.2364	0.4391	0.1307	0.0537	0.3165	0.3060	0.4676
26	0.3091	0.5297	0.2219	0.0512	0.3165	0.2760	0.3462

NCNji							
Sample No.	**pH**	**TDS (mg/l)**	**Turbidity (NTU)**	**Fe (mg/l)**	**Mn (mg/l)**	**Total hardness (mg/l)**	**Cl (mg/l)**
27	0.2000	0.5297	0.2827	0.0409	0.3165	0.3224	0.3178
28	0.8636	0.3988	0.0866	0.0179	0.3165	0.2951	0.3745
29	0.6545	0.7466	0.5182	0.0256	0.3165	0.3579	0.4555
30	0.4182	0.6591	0.3541	0.0460	0.3165	0.2459	0.3381
31	0.4091	0.9825	0.4544	0.0665	0.0633	0.4262	0.3259
32	0.6000	0.0974	0.2356	0.0512	0.2152	0.7541	0.0646
33	0.8455	0.2230	0.1185	0.0358	0.1139	0.4262	0.2085
34	0.4455	0.3508	0.1049	0.0409	0.0127	0.5082	0.4676
35	0.2182	0.3097	0.2234	0.0588	0.4557	0.3989	0.3522
36	0.2636	0.2610	0.0167	0.0486	0.0253	0.5082	0.3947
37	0.1909	0.2770	0.0000	0.0307	0.2025	0.5628	0.3947
38	0.2636	0.4802	0.0790	0.0639	0.2658	0.6995	0.6113
39	0.1364	0.6027	0.0942	0.0460	0.0000	0.8907	0.6802
40	0.2182	0.7671	0.1018	0.0665	0.4304	0.9454	1.0000
41	0.2364	0.4688	0.0106	0.0435	0.0127	0.5628	0.7976
42	0.0182	0.5761	0.1581	0.0614	1.0000	1.0000	0.6215
43	0.4727	0.3592	0.3450	1.0000	0.1392	0.4262	0.4090
44	0.3545	0.3569	0.1125	0.0716	0.3165	0.2459	0.2996
45	0.2273	0.5327	0.2036	0.0537	0.3165	0.3579	0.4028
46	0.3273	0.3242	0.1094	0.0256	0.3165	0.2869	0.4433
47	0.9091	0.8813	0.5243	0.0921	0.3165	0.4536	0.6255
48	0.1091	0.4597	0.1383	0.0512	0.3165	0.3279	0.4717
49	0.6455	0.3714	0.1398	0.2327	0.3165	0.3661	0.5567
50	0.3000	0.1895	0.1018	0.1049	0.1101	0.4481	0.2166
51	0.3182	0.1659	0.0775	0.0895	0.1949	0.1858	0.3745
52	0.2364	0.2679	0.1261	0.1023	0.1899	0.3169	0.2947
53	0.3000	0.5685	0.0988	0.1560	0.2025	0.3716	0.5247
54	0.3909	0.1537	0.1216	0.1483	0.1899	0.3279	0.2947
55	0.4727	0.3455	0.0714	0.0870	0.1646	0.2732	0.3235
56	0.8909	0.2549	0.0821	0.1072	0.0354	0.1967	0.2832
57	0.5545	0.1545	0.0638	0.1069	0.0709	0.3279	0.2430
58	0.7273	0.1469	0.0593	0.1448	0.0215	0.1639	0.3292
59	0.8727	0.4909	0.1337	0.1895	0.1684	0.5246	0.3005
60	0.6364	0.4079	0.1337	0.1558	0.1646	0.3934	0.2717
61	0.5364	0.3858	0.0669	0.0877	0.0924	0.3060	0.2775
62	0.5636	0.2268	0.1033	0.1072	0.0633	0.2077	0.2372
63	0.6091	0.2770	0.0851	0.1317	0.0443	0.3934	0.2545
64	0.6727	0.1317	0.0927	0.0977	0.0709	0.1311	0.3062

NCNji							
Sample No.	**pH**	**TDS (mg/l)**	**Turbidity (NTU)**	**Fe (mg/l)**	**Mn (mg/l)**	**Total hardness (mg/l)**	**Cl (mg/l)**
65	0.1636	0.1195	0.1337	0.1358	0.0785	0.2514	0.1683
66	0.2636	0.3067	0.1216	0.1292	0.0570	0.3169	0.2257
67	0.5273	0.1773	0.1155	0.0854	0.0342	0.1639	0.3810
68	0.6182	0.1697	0.1368	0.0708	0.3189	0.6230	0.2224
69	0.6455	0.5563	0.1109	0.0077	0.3165	0.1639	0.6245
70	0.6273	0.2199	0.1444	0.0228	0.0797	0.1202	0.2028
71	0.8000	0.2093	0.1246	0.0182	0.0481	0.0874	0.1510
72	0.8000	0.1362	0.0760	0.0199	0.0519	0.0656	0.1338
73	0.7000	0.1248	0.0942	0.0156	0.0494	0.0546	0.1280
74	0.7455	0.1286	0.0684	0.0028	0.0165	0.0656	0.1162
75	0.7818	0.2291	0.0927	0.0115	0.0785	0.0874	0.1854
76	0.4545	0.5312	0.1109	0.0537	0.0253	0.1421	0.9131
77	0.4000	0.2062	0.2523	0.0435	0.0253	0.0109	0.3619
78	0.5273	0.2763	0.0380	0.0281	0.0253	0.0546	0.4722
79	0.5727	0.3721	0.1991	0.0460	0.0380	0.0000	0.5273
80	0.2818	0.0129	0.0684	0.2379	0.0506	0.2049	0.0972
81	0.4818	0.0167	0.0426	0.0102	0.0633	0.2186	0.0526
82	0.6091	0.0457	0.0623	0.0435	0.0759	0.3060	0.1964
83	0.6909	0.0236	0.1201	0.0128	0.1646	0.2022	0.0931
84	0.7273	0.0266	0.0517	0.0179	0.0380	0.1093	0.0486
85	0.6455	0.0030	0.0243	0.0077	0.2152	0.0984	0.0607
86	0.7000	0.0304	0.0015	0.0000	0.0506	0.1093	0.0405
87	0.5182	0.0000	0.0638	0.0639	0.0506	0.0929	0.0344
88	0.7000	0.0297	0.0426	0.0051	0.0127	0.1530	0.1134
89	0.3909	0.0076	0.0426	0.0153	0.0886	0.2732	0.1215
90	0.2636	0.0441	0.1049	0.0051	0.0506	0.2732	0.1215
91	0.4182	0.0647	0.3024	0.0870	0.0886	0.2077	0.1275
92	0.4455	0.0411	0.2112	0.0179	0.0253	0.1093	0.0729
93	0.3909	0.8311	0.4726	0.0813	0.0000	0.2650	0.6978
94	0.4818	0.5266	0.3283	0.3026	0.0000	0.2923	0.7054
95	0.3818	0.8006	0.4833	0.0977	0.0000	0.2322	0.7046
96	0.2000	0.5723	1.0000	0.2954	0.0000	0.3415	0.5571
97	0.5000	0.1142	0.7462	0.1074	0.0633	0.4918	0.1397
98	0.4091	0.0396	0.4347	0.0256	0.3418	0.3388	0.0000
99	0.2455	0.1811	0.1824	0.0793	0.0633	0.5355	0.3462
100	0.3455	0.1766	0.0912	0.0563	0.2025	0.3716	0.1599

According to (Gorgij et al. 2017), entropy weighting can enhance the WQI model's ability to measure water quality. Because the E_NWQI approach minimizes subjective judgments regarding parameter weighting and assists decision makers in assessing groundwater quality, it offers a more practical and accurate water quality assessment.

The entropy weight (E_NW_i) and information entropy (IEi) of all the analyzed parameters are shown in Table 5.5. The value of E_NWQI is calculated by multiplying entropy weightage with the quality rating factor However, the final EWQI and the ranking of the water samples are presented in Table 5.6.

Table 5.5. Values of Entropy based weights (E_NW_i) and Information Entropy (IEi) of individual parameters.

WQP	pH	TDS (mg/l)	Turbidity (NTU)	Iron (mg/l)	Manganese (mg/l)	Total Hardness (mg/l)	Chloride (mg/l)
IEi	0.9972	0.9966	0.9979	0.9991	0.9983	0.9980	0.9977
E_NW_i	0.1842	0.2240	0.1391	0.0617	0.1135	0.1294	0.1481

Table 5.6. Values of E_NWQI.

Sample Number	E_NWQI	Rank	Sample Number	E_NWQI	Rank	Sample Number	E_NWQI	Rank
1	103.0518	3	2	147.1067	3	3	107.1614	3
4	130.0489	3	5	99.6439	2	6	204.7641	4
7	117.9712	3	8	163.3284	3	9	111.9625	3
10	122.1335	3	11	108.2376	3	12	195.2678	3
13	138.9862	3	14	115.9142	3	15	109.4965	3
16	114.5162	3	17	112.9874	3	18	120.4254	3
19	163.2156	3	20	111.8171	3	21	99.6986	2
22	101.237	3	23	213.9703	4	24	125.4504	3
25	118.0574	3	26	128.2433	3	27	131.7802	3
28	114.263	3	29	175.8784	3	30	148.0713	3
31	159.0204	3	32	101.9795	3	33	88.1106	2
34	89.8334	2	35	130.4015	3	36	73.6451	2
37	87.8240	2	38	125.9223	3	39	114.0689	3
40	176.4464	3	41	96.4864	2	42	208.5417	4
43	197.4545	3	44	108.2458	3	45	129.454	3
46	107.1484	3	47	200.3889	4	48	118.6735	3

Sample Number	E_NWQI	Rank	Sample Number	E_NWQI	Rank	Sample Number	E_NWQI	Rank
49	138.8511	3	50	83.2054	2	51	84.6183	2
52	95.3125	2	53	124.8419	3	54	94.2219	2
55	94.3712	2	56	82.7165	2	57	75.6768	2
58	74.4003	2	59	127.8985	3	60	112.8203	3
61	90.2093	2	62	79.9951	2	63	87.0731	2
64	75.0223	2	65	73.7367	2	66	85.7675	2
67	76.5153	2	68	109.8378	3	69	126.201	3
70	75.8178	2	71	70.2174	2	72	60.9288	2
73	59.5708	2	74	53.9782	2	75	71.4139	2
76	107.6569	3	77	80.6662	2	78	69.9288	2
79	93.8200	2	80	65.6061	2	81	48.0276	1
82	63.3413	2	83	68.4319	2	84	48.3994	1
85	58.5715	2	86	43.1257	1	87	49.1519	1
88	47.0081	1	89	52.2518	2	90	54.1557	2
91	84.1391	2	92	59.6193	2	93	154.1064	3
94	142.8551	3	95	153.9024	3	96	199.4464	3
97	135.2192	3	98	111.7832	3	99	88.9503	2
100	82.9961	2						

As shown in Table 5.6, computed WQI values using various weight techniques have been divided into five divisions. MS spreadsheet programs are used to determine the WQI values at various places using the aforesaid weight techniques. Out of 100 samples, almost 50% of samples are classified as Poor, only 4% samples are very poor, 41% samples are classified as Good for drinking and only 5% samples are of excellent quality for drinking as mentioned in Table 4.1. There are no samples which are unsuitable for drinking.

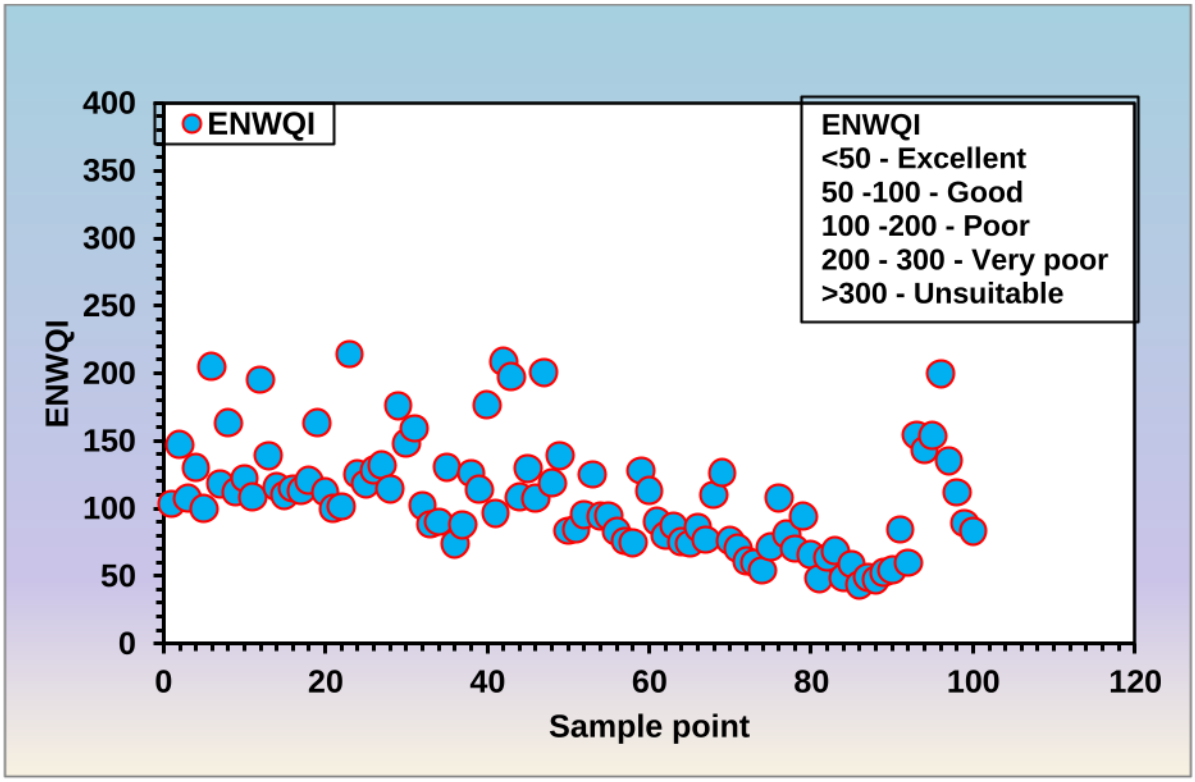


Fig. 5.1. Graphical representation of Entropy WQI against sample points.

5.1.2.2. Computation of CRITIC Weight WQI (C_WWQI)

The WQI technique, which is based on CRITIC, is used to evaluate groundwater quality while taking into account 100 samples. The CRITIC WQI indicates excellent to very poor class for drinking, with a range of 41.08 to 222.16.

Table 5.7. Values of correlation coefficients

(CRji)							
	pH	TDS	Turbidity	Iron	Manganese	Total Hardness	Chloride
pH	1.0000	-0.1720	-0.0680	-0.0260	-0.2500	-0.3240	-0.2560
TDS	-0.1720	1.0000	0.5730	0.0740	0.3760	0.3580	0.7380
Turbidity	-0.0680	0.5730	1.0000	0.2260	0.1240	0.1390	0.2680
Iron	-0.0260	0.0740	0.2260	1.0000	-0.1180	0.1410	0.1510
Manganese	-0.2500	0.3760	0.1240	-0.1180	1.0000	0.4070	0.2300
Hardness	-0.3240	0.3580	0.1390	0.1410	0.4070	1.0000	0.4240
Chloride	-0.2560	0.7380	0.2680	0.1510	0.2300	0.4240	1.0000

The standard deviation (SDVi), information amount (IFi) and CRITIC weight of individual parameters (CWp) are shown in Table 5.8. The value of C_WWQI is calculated by multiplying CRITIC weight with the quality rating factor. However, the final C_WWQI and the ranking of the water samples are presented in Table 5.9.

Table 5.8. Values of standard deviation, information amount and weight of individual parameters.

WQP	SDVi	IFi	CWp
pH	0.2362	1.6761	0.2515
TDS(mg/l)	0.2445	0.9908	0.1487
Turbidity (NTU)	0.1741	0.8251	0.1238
Fe (mg/l)	0.1094	0.6071	0.0911
Mn (mg/l)	0.1542	0.8065	0.1210
TH (mg/l)	0.1820	0.8837	0.1326
Cl (mg/l)	0.1970	0.8757	0.1314

Table 5.9. Values of C_WWQI.

Sample number	CWWQI	Rank	Sample number	CWWQI	Rank	Sample number	CWWQI	Rank
1	95.4282	2	2	127.2055	3	3	100.7244	3
4	124.2027	3	5	91.2675	2	6	177.5717	3
7	106.0514	3	8	143.2571	3	9	101.2194	3
10	112.6804	3	11	97.7529	2	12	173.6433	3
13	124.4459	3	14	103.3838	3	15	97.5331	2
16	102.9389	3	17	100.9019	3	18	109.3191	3
19	148.1716	3	20	99.5477	2	21	92.4314	2
22	94.9234	2	23	189.5238	3	24	119.2319	3
25	106.6692	3	26	114.7899	3	27	116.9068	3
28	106.1374	3	29	155.5647	3	30	131.0795	3
31	134.0051	3	32	98.9922	2	33	83.1705	2
34	79.3169	2	35	122.0081	3	36	65.5046	2

Sample number	CWWQI	Rank	Sample number	CWWQI	Rank	Sample number	CWWQI	Rank
37	79.5679	2	38	114.2205	3	39	96.7879	2
40	158.5398	3	41	82.5899	2	42	197.3151	3
43	222.1612	4	44	100.457	3	45	115.6785	3
46	97.6885	2	47	180.6458	3	48	106.0757	3
49	137.3354	3	50	79.0174	2	51	80.4712	2
52	89.0534	2	53	114.6248	3	54	92.8030	2
55	87.4316	2	56	79.6266	2	57	73.5089	2
58	74.0611	2	59	123.6032	3	60	107.7043	3
61	82.6045	2	62	75.9578	2	63	83.3283	2
64	72.9241	2	65	70.9877	2	66	79.3922	2
67	71.6516	2	68	107.2705	3	69	112.3427	3
70	68.7425	2	71	64.1915	2	72	56.9733	2
73	54.9962	2	74	49.1723	1	75	65.0377	2
76	92.4360	2	77	71.5429	2	78	61.2708	2
79	82.4178	2	80	70.1616	2	81	45.2558	1
82	61.3522	2	83	66.3307	2	84	46.6454	1
85	57.7636	2	86	41.0802	1	87	48.6492	1
88	44.3227	1	89	49.3628	1	90	48.6737	1
91	80.3332	2	92	54.4325	2	93	130.6931	3
94	135.8411	3	95	131.5266	3	96	183.6386	3
97	127.1422	3	98	106.7409	3	99	82.2170	2
100	78.2206	2						

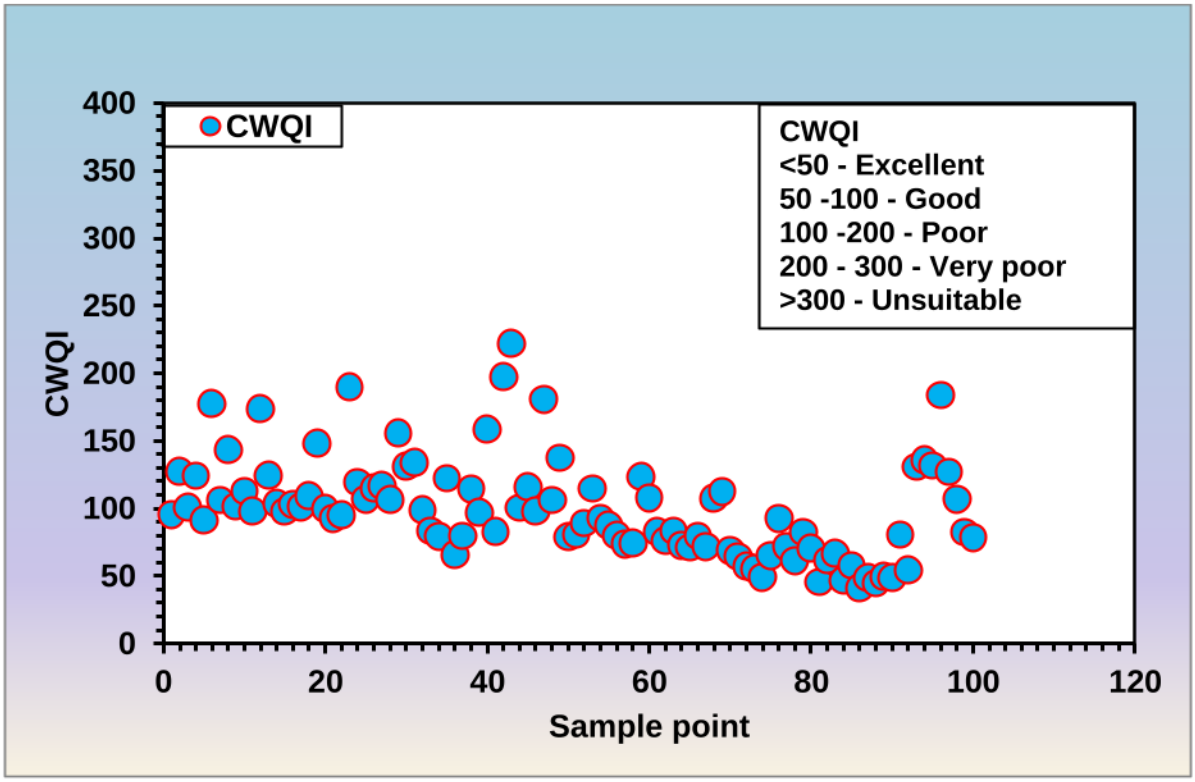


Fig. 5.2. Graphical representation of CRITIC WQI against sample points.

Water quality includes excellent (8% samples), good (47% samples), poor (44% samples) and very poor (1% samples) from 100 samples as per above mentioned Table 4.1. There are no samples which are unsuitable for drinking.

The CRITIC weighting method uses trade-offs between several variables and the STDV concept to determine WQI. It also restricts individual opinion when determining how much weight to give each characteristic. Thus, the WQI system's objective weighting, as opposed to subjective weighting, can increase the precision of water quality evaluation and model predictions. We have taken into account a restricted number of WQPs in WQI calculation due to the dataset's availability. Different scenarios for regional water quality can be provided and the process of assessing water quality can be improved by including a greater number of WQPs in the WQI estimation (Das and Das, 2024). Microsoft Excel 2013 is used to construct the Entropy and CRITIC based WQI. The WQI model is developed utilizing these WQI scores.

5.1.2.3. Comparison of CRITIC and Entropy Method Through Eclipsing and Ambiguity

In the CRITIC method, eight sampling points in which all the considered water quality parameters are within the permissible limit for drinking and after observing their respective final WQI scores which suggest water is acceptable in both WQI methods in those sampling pts. Thus both methods are free from ambiguity.

The eclipsing issue has been spotted/ identified at 74 sampling points (over-estimate eclipsing 71 sampling points and under-estimate eclipsing in 3 sampling points) for the Entropy WQI method. The eclipsing issue has been spotted at 77 sampling points (over estimate eclipsing at 76 sampling points and under estimate eclipsing at 1 sampling point) for the CRITIC WQI method. Therefore, the Entropy WQI method has fewer/ less eclipsing issues as compared to the Critic WQI.

5.1.3. MLR Model Preparation

5.1.3.1. Input Combination for Model Preparation

Both the outcome of this index and all factors associated with the original WQI were assessed for their sample distribution. The majority of variables do not follow a normal distribution, according to the K–S and S–W tests, which produced a p value < 0.05 (p-value provided by SPSS, quoted under Sig. for Kolmogorov-Smirnov and Shapiro-Wilk). As a result, Spearman's coefficient will be the correlation coefficient employed in the study's subsequent phase.

Table 5.10. Values of significance in K-S and S-W Tests.

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
pH	0.071	100	0.200*	0.981	100	0.151
TDS	0.080	100	0.110	0.949	100	0.001
Turbidity	0.222	100	0.000	0.771	100	0.000
Fe	0.256	100	0.000	0.456	100	0.000
Mn	0.206	100	0.000	0.803	100	0.000
TH	0.124	100	0.001	0.902	100	0.000
Cl	0.102	100	0.013	0.958	100	0.003
WQI	0.080	100	0.115	0.949	100	0.001

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

The Kolmogorov Smirnov test produces test statistics that are used (along with a degree of freedom parameter) to test for normality. Here it can be seen that the Kolmogorov Smirnov statistic takes various value which has degrees of freedom which equals the

number of data points, i.e. 100. Table 5.10 shows the specific test results including the Shapiro-Wilk Test Statistic (Statistic), the degrees of freedom (df) the two-tailed significance or p-value (Sig.) also. The Shapiro–Wilk test is a more appropriate method for small sample sizes (<50 samples) although it can also handle larger sample sizes while the Kolmogorov–Smirnov test is used for $n \geq 50$.

In the Spearman correlation matrix, the correlation coefficient is low in the case of pH and WQI. So, one input combination is TDS, Turbidity, Fe, Mn, Total Hardness and Chloride with WQI value.

In the Spearman correlation matrix, it is found that pH and TDS are negatively correlated. In coastal areas (Das and Das, 2024) found that TDS and Chloride are highly correlated and in our study, TDS and Chloride are highly correlated but the correlation coefficient of chloride with WQI is less as compared to TDS and the variance of chloride is also less as compared to TDS. For determining the second input combination, chloride is not considered. So, another input combination is pH, TDS, Turbidity, Fe, Mn, and Total Hardness with WQI value.

Considering the above two input conditions, in the final combination pH and Chloride both are not considered. So, the final combination is TDS, turbidity, Fe, Mn, and total hardness with WQI value.

5.1.3.2. Dataset Distribution for model preparation

Table 5.11. Correlation coefficient of Spearman Correlation Matrix.

	pH	TDS	Turbidity	Fe	Mn	TH	Cl	WQI
pH	1.000	-0.246*	-0.146	-0.129	-0.196	-0.322**	-0.276**	-0.225*
TDS	-0.246*	1.000	0.540**	0.217*	0.366**	0.396**	0.787**	0.868**
Turbidity	-0.146	0.540**	1.000	0.218*	0.358**	0.228*	0.307**	0.748**
Fe	-0.129	0.217*	0.218*	1.000	-0.094	0.388**	0.287**	0.337**
Mn	-0.196	0.366**	0.358**	-0.094	1.000	0.324**	0.172	0.553**
TH	-0.322**	0.396**	0.228*	0.388**	0.324**	1.000	0.384**	0.539**
Cl	-0.276**	0.787**	0.307**	0.287**	0.172	0.384**	1.000	0.688**
WQI	-0.225*	0.868**	0.748**	0.337**	0.553**	0.539**	0.688**	1.000
*. Correlation is significant at the 0.05 level (2-tailed).								
**. Correlation is significant at the 0.01 level (2-tailed).								

From various studies, it is seen that researchers assumed different distribution ratios in their studies and there are no specific rules for dataset ratio distribution. So four dataset distribution ratios are taken into consideration, taking into account the input parameters and the three combinations indicated above. The four ratios are 70:30, 75:25, 80:20 and 85:15.

5.1.3.3. Identification of MLR Based New G_{WQI} Models

In order to create new GWQI equations, a stepwise MLR technique was developed in this study. This MLR analysis yielded twelve models. Table 5.12 displays the findings of the analysis. New GWQI equations from this study are represented by the tables mentioned below. Due to the unequal relevance of the model variables, considering the value of adjusted R^2, twelve suitable regression models were taken into consideration (Table 5.12). As a measure of the strength of the relationship, coefficients of determination (R^2) for multiple linear regressions (MLR) were also more quickly explained than coefficients of multiple correlations. This is because multiple R^2 represents the percentage of the total variance in the presence of a parameter that may be related to the predictor parameter outcomes (Das et al. 2023). The independent parameters of 12 models are significant (significance level < 0.05).

Table 5.12. Model Summary.

Model summary of 70:30 model when pH is not considered					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Independent Variables
1	0.865	0.749	0.745	21.5046	Predictors: (Constant), TDS
2	0.915	0.838	0.833	17.4239	Predictors: (Constant), TDS, Turbidity
3	0.955	0.913	0.909	12.8561	Predictors: (Constant), TDS, Turbidity, Mn
4	0.991	0.983	0.982	5.7707	Predictors: (Constant), TDS, Turbidity, Mn, Fe
5	0.995[e]	0.991	0.990	4.2552	Predictors: (Constant), TDS, Turbidity, Mn, Fe, Cl
6	0.998	0.995	0.995	3.0795	Predictors: (Constant), TDS, Turbidity, Mn, Fe, Cl, TH
MODEL 1: Considering model 6, G_{WQI} **= 9.414 + 0.047 TDS+ 13.647 Tur + 111.032 Mn + 20.699 Fe + 0.051 Cl+ 0.054 TH ± 1.252**					

Model summary of 75:25 model when pH is not considered					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Independent Variables
1	0.850	0.723	0.719	22.6955	Predictors: (Constant), TDS
2	0.910	0.829	0.824	17.9615	Predictors: (Constant), TDS, Turbidity
3	0.960	0.922	0.919	12.2175	Predictors: (Constant), TDS, Turbidity, Mn
4	0.991	0.981	0.980	5.9963	Predictors: (Constant), TDS, Turbidity, Mn, Fe
5	0.996	0.991	0.991	4.1546	Predictors: (Constant), TDS, Turbidity, Mn, Fe, TH
6	0.998	0.996	0.995	2.9862	Predictors: (Constant), TDS, Turbidity, Mn, Fe, TH, Cl
MODEL 2: Considering model 6, G_{WQI} **= 10.741 + 0.046 TDS + 13.864 Tur + 109.858 Mn + 20.334 Fe + 0.050 Cl + 0.057 TH ± 1.068**					
Model summary of 80:20 model when pH is not considered					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Independent Variables
1	0.853	0.728	0.725	20.0459	Predictors: (Constant), TDS
2	0.922	0.851	0.847	14.9424	Predictors: (Constant), TDS, Turbidity
3	0.951	0.904	0.900	12.0624	Predictors: (Constant), TDS, Turbidity, Fe
4	0.988	0.977	0.975	5.9906	Predictors: (Constant), TDS, Turbidity, Fe, Mn
5	0.994	0.988	0.987	4.3907	Predictors: (Constant), TDS, Turbidity, Fe, Mn, Cl
6	0.997	0.994	0.994	3.0654	Predictors: (Constant), TDS, Turbidity, Fe, Mn, Cl, TH
MODEL 3 : Considering model 6, G_{WQI} **= 10.027 + 0.047 TDS + 13.907 Tur + 109.240 Mn + 20.889 Fe + 0.051 Cl + 0.053 TH ± 1.094**					

Model summary of 85:15 model when pH is not considered					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Independent Variables
1	0.882	0.777	0.774	19.8202	Predictors: (Constant), TDS
2	0.931	0.867	0.863	15.4153	Predictors: (Constant), TDS, Mn
3	0.973	0.947	0.945	9.7579	Predictors: (Constant), TDS, Mn, Fe
4	0.991	0.982	0.981	5.7519	Predictors: (Constant), TDS, Mn, Fe, Turbidity
5	0.995	0.990	0.989	4.2847	Predictors: (Constant), TDS, Mn, Fe, Turbidity, Cl
6	0.997	0.995	0.994	3.1175	Predictors: (Constant), TDS, Mn, Fe, Turbidity, Cl, TH
MODEL 4: Considering model 6, **G_{WQI} = 9.485 + 0.047 TDS + 13.805 Tur + 108.882 Mn + 20.601 Fe + 0.051 Cl + 0.057 TH ± 1.091**					
Model summary of 70:30 model when Cl is not considered					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Independent Variables
1	0.865	0.749	0.745	21.5046	Predictors: (Constant), TDS
2	0.915	0.838	0.833	17.4239	Predictors: (Constant), TDS, Turbidity
3	0.955	0.913	0.909	12.8561	Predictors: (Constant), TDS, Turbidity, Mn
4	0.991	0.983	0.982	5.7707	Predictors: (Constant), TDS, Turbidity, Mn, Fe
5	0.995	0.990	0.989	4.3926	Predictors: (Constant), TDS, Turbidity, Mn, Fe, TH
6	0.997	0.993	0.993	3.6347	Predictors: (Constant), TDS, Turbidity, Mn, Fe, TH, pH
MODEL 5: Considering model 6, **G_{WQI} = -68.815 + 0.061TDS + 12.596Tur + 102.876Mn + 21.743Fe + 0.078TH + 9.950 pH ± 14.096**					

Model summary of 75:25 model when Cl is not considered					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Independent Variables
1	0.850	0.723	0.719	22.6955	Predictors: (Constant), TDS
2	0.910	0.829	0.824	17.9615	Predictors: (Constant), TDS, Turbidity
3	0.960	0.922	0.919	12.2175	Predictors: (Constant), TDS, Turbidity, Mn
4	0.991	0.981	0.980	5.9963	Predictors: (Constant), TDS, Turbidity, Mn, Fe
5	0.996	0.991	0.991	4.1546	Predictors: (Constant), TDS, Turbidity, Mn, Fe, TH
6	0.997	0.994	0.994	3.4349	Predictors: (Constant), TDS, Turbidity, Mn, Fe, TH, pH
MODEL 6 : Considering model 6, G_{WQI} **= -65.879 + 0.058TDS + 12.842Tur + 111.074Mn + 21.272Fe + 0.079TH + 9.662 pH ± 13.325**					
Model summary of 80:20 model when Cl is not considered					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Independent Variables
1	0.882	0.777	0.774	19.8202	Predictors: (Constant), TDS
2	0.931	0.867	0.863	15.4153	Predictors: (Constant), TDS, Mn
3	0.973	0.947	0.945	9.7579	Predictors: (Constant), TDS, Mn, Fe
4	0.991	0.982	0.981	5.7519	Predictors: (Constant), TDS, Mn, Fe, Turbidity
5	0.995	0.990	0.989	4.4148	Predictors: (Constant), TDS, Mn, Fe, Turbidity, TH
6	0.997	0.993	0.993	3.6098	Predictors: (Constant), TDS, Mn, Fe, Turbidity, TH, pH
MODEL 7: Considering model 6, G_{WQI} **= -69.932 + 0.062 TDS + 11.698 Tur + 108.966 Mn + 21.801 Fe + 0.076 TH + 10.096 pH ± 12.936**					

Model summary of 85:15 model when Cl is not considered					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Independent Variables
1	0.853	0.728	0.725	20.0459	Predictors: (Constant), TDS
2	0.922	0.851	0.847	14.9424	Predictors: (Constant), TDS, Turbidity
3	0.951	0.904	0.900	12.0624	Predictors: (Constant), TDS, Turbidity, Fe
4	0.988	0.977	0.975	5.9905	Predictors: (Constant), TDS, Turbidity, Fe, Mn
5	0.994	0.987	0.987	4.4126	Predictors: (Constant), TDS, Turbidity, Fe, Mn, TH
6	0.996	0.991	0.991	3.6981	Predictors: (Constant), TDS, Turbidity, Fe, Mn, TH, pH
MODEL 8 : Considering model 6, G_{WQI} **= -66.581 + 0.060TDS + 12.664Tur + 104.948Mn + 21.872Fe + 0.080TH + 9.699 pH ± 12.920**					
Model summary of 70:30 model when pH and Cl both are not considered					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Independent Variables
1	0.865	0.749	0.745	21.5046	Predictors: (Constant), TDS
2	0.915	0.838	0.833	17.4239	Predictors: (Constant), TDS, Turbidity
3	0.955	0.913	0.909	12.8561	Predictors: (Constant), TDS, Turbidity, Mn
4	0.991	0.983	0.982	5.7707	Predictors: (Constant), TDS, Turbidity, Mn, Fe
5	0.995	0.990	0.989	4.3926	Predictors: (Constant), TDS, Turbidity, Mn, Fe, TH
MODEL 9: Considering model 5, G_{WQI} **= 8.572 + 0.062TDS + 12.529Tur + 102.001Mn + 21.726Fe + 0.068TH ± 1.780**					

Model summary of 75:25 model when pH and Cl both are not considered					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Independent Variables
1	0.850	0.723	0.719	22.6955	Predictors: (Constant), TDS
2	0.910	0.829	0.824	17.9615	Predictors: (Constant), TDS, Turbidity
3	0.960	0.922	0.919	12.2175	Predictors: (Constant), TDS, Turbidity, Mn
4	0.991	0.981	0.980	5.9963	Predictors: (Constant), TDS, Turbidity, Mn, Fe
5	0.996	0.991	0.991	4.1546	Predictors: (Constant), TDS, Turbidity, Mn, Fe, TH
MODEL 10: Considering model 5, G_{WQI} **= 10.277 + 0.057TDS + 12.948Tur + 108.401Mn + 20.996Fe + 0.071TH ± 1.484**					
Model summary of 80:20 model when pH and Cl both are not considered					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Independent Variables
1	0.882	0.777	0.774	19.8202	a. Predictors: (Constant), TDS
2	0.931	0.867	0.863	15.4153	b. Predictors: (Constant), TDS, Mn
3	0.973	0.947	0.945	9.7579	c. Predictors: (Constant), TDS, Mn, Fe
4	0.991	0.982	0.981	5.7519	d. Predictors: (Constant), TDS, Mn, Fe, Turbidity
5	0.995	0.990	0.989	4.4147	e. Predictors: (Constant), TDS, Mn, Fe, Turbidity, TH
MODEL 11: Considering model 5, G_{WQI} **= 9.098 + 0.061 TDS + 11.946 Tur + 105.860 Mn + 21.940 Fe + 0.065 TH ± 1.541**					
Model summary of 85:15 model when pH and Cl both are not considered					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Independent Variables
1	0.853	0.728	0.725	20.0459	Predictors: (Constant), TDS
2	0.922	0.851	0.847	14.9424	Predictors: (Constant), TDS, Turbidity
3	0.951	0.904	0.900	12.0624	Predictors: (Constant), TDS, Turbidity, Fe

4	0.988	0.977	0.975	5.9905	Predictors: (Constant), TDS, Turbidity, Fe, Mn
5	0.994	0.987	0.987	4.4126	Predictors: (Constant), TDS, Turbidity, Fe, Mn, TH
MODEL 12: Considering model 5, G_{WQI} **= 8.885 + 0.060 TDS + 12.728 Tur + 102.168 Mn + 21.752 Fe + 0.071 TH ± 1.567**					

Table 5.13. Model wise coefficient values.

Model No.	Ratio	Coefficient					
		Parameter	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
			B	Std. Error	Beta		
Model 1	70:30	(Constant)	9.414	1.252		7.521	0.0000
		TDS	0.047	0.002	0.362	19.981	0.0000
		Turbidity	13.647	0.435	0.383	31.359	0.0000
		Mn	111.032	3.857	0.326	28.784	0.0000
		Fe	20.699	0.81	0.243	25.543	0.0000
		Cl	0.051	0.006	0.117	8.199	0.0000
		TH	0.054	0.007	0.082	7.694	0.0000
Model 2	75:25	(Constant)	10.741	1.068		10.054	0.0000
		TDS	0.046	0.002	0.356	22.517	0.0000
		Turbidity	13.864	0.367	0.407	37.809	0.0000
		Mn	109.858	3.167	0.33	34.685	0.0000
		Fe	20.334	0.787	0.224	25.85	0.0000
		TH	0.057	0.006	0.095	9.405	0.0000
		Cl	0.05	0.006	0.111	8.097	0.0000
Model 3	80:20	(Constant)	10.027	1.094		9.169	0.0000
		TDS	0.047	0.002	0.38	20.032	0.0000
		Mn	109.24	3.226	0.334	33.866	0.0000
		Fe	20.889	0.822	0.227	25.412	0.0000

Model No.	Ratio	Coefficient					
		Parameter	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
			B	Std. Error	Beta		
		Turbidity	13.907	0.548	0.338	25.378	0.0000
		Cl	0.051	0.006	0.125	8.683	0.0000
		TH	0.053	0.006	0.085	8.172	0.0000
Model 4	85:15	(Constant)	9.485	1.091		8.696	0.0000
		TDS	0.047	0.002	0.392	23.806	0.0000
		Turbidity	13.805	0.375	0.405	36.786	0.0000
		Fe	20.601	0.781	0.248	26.395	0.0000
		Mn	108.882	3.563	0.3	30.557	0.0000
		Cl	0.051	0.005	0.132	9.257	0.0000
		TH	0.057	0.006	0.092	9.169	0.0000
Model 5	70:30	(Constant)	-68.815	14.096		-4.882	0.0000
		TDS	0.061	0.002	0.47	32.366	0.0000
		Turbidity	12.596	0.488	0.353	25.814	0.0000
		Mn	102.876	4.366	0.302	23.562	0.0000
		Fe	21.743	0.945	0.255	23.01	0.0000
		TH	0.078	0.008	0.118	9.421	0.0000
		pH	9.95	1.803	0.059	5.52	0.0000
Model 6	75:25	(Constant)	-65.879	13.325		-4.944	0.0000
		TDS	0.058	0.002	0.456	35.399	0.0000
		Turbidity	12.842	0.402	0.377	31.974	0.0000
		Mn	111.074	3.667	0.333	30.289	0.0000
		Fe	21.272	0.901	0.234	23.604	0.0000
		TH	0.079	0.007	0.131	11.557	0.0000
		pH	9.662	1.683	0.058	5.739	0.0000
Model 7	80:20	(Constant)	-69.932	12.936		-5.406	0.0000

Model No.	Ratio	Coefficient					
		Parameter	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
			B	Std. Error	Beta		
		TDS	0.062	0.002	0.503	33.362	0.0000
		Mn	108.966	3.742	0.333	29.119	0.0000
		Fe	21.801	0.942	0.237	23.15	0.0000
		Turbidity	11.698	0.58	0.285	20.183	0.0000
		TH	0.076	0.008	0.124	10.184	0.0000
		pH	10.096	1.645	0.065	6.138	0.0000
Model 8	85:15	(Constant)	-66.581	12.92		-5.153	0.0000
		TDS	0.06	0.002	0.501	36.19	0.0000
		Turbidity	12.664	0.431	0.372	29.41	0.0000
		Fe	21.872	0.93	0.264	23.523	0.0000
		Mn	104.948	4.235	0.289	24.78	0.0000
		TH	0.08	0.007	0.129	10.802	0.0000
		pH	9.699	1.652	0.064	5.871	0.0000
Model 9	70:30	(Constant)	8.572	1.78		4.817	0.0000
		TDS	0.062	0.002	0.471	26.88	0.0000
		Turbidity	12.529	0.589	0.351	21.254	0.0000
		Mn	102.001	5.273	0.3	19.344	0.0000
		Fe	21.726	1.142	0.255	19.025	0.0000
		TH	0.068	0.01	0.103	6.941	0.0000
Model 10	75:25	(Constant)	10.277	1.484		6.924	0.0000
		TDS	0.057	0.002	0.448	28.912	0.0000
		Turbidity	12.948	0.485	0.38	26.682	0.0000
		Mn	108.401	4.4	0.325	24.639	0.0000
		Fe	20.996	1.088	0.231	19.29	0.0000
		TH	0.071	0.008	0.118	8.765	0.0000

Model No.	Ratio	Coefficient					
		Parameter	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
			B	Std. Error	Beta		
Model 11	80:20	(Constant)	9.098	1.541		5.903	0.0000
		TDS	0.061	0.002	0.5	27.123	0.0000
		Mn	105.86	4.535	0.323	23.345	0.0000
		Fe	21.94	1.151	0.238	19.056	0.0000
		Turbidity	11.946	0.707	0.291	16.895	0.0000
		TH	0.065	0.009	0.105	7.302	0.0000
Model 12	85:15	(Constant)	8.885	1.567		5.669	0.0000
		TDS	0.06	0.002	0.501	30.297	0.0000
		Turbidity	12.728	0.514	0.374	24.782	0.0000
		Fe	21.752	1.109	0.262	19.612	0.0000
		Mn	102.168	5.022	0.282	20.345	0.0000
		TH	0.071	0.009	0.116	8.273	0.0000

In this study, it is seen that the significance value of all 12 models is zero.

5.1.3.4. Validation of Models by T-test

Student t-test has been used to compare the 30% (for 70:30 ratio), 25% (for 75:25 ratio), 20% (for 80:20 ratio), and 15% (for 85:15 ratio) sample's original and updated GWQI in order to validate the suggested models. There is no discernible difference between the original GWQI's results and those from the new one. As the significance is higher than 0.05, the t-test result validates the efficacy of the new GWQI equations.

Table 5.14. Model wise *t*-test result**.**

Paired Samples Test									
		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
MODEL 1									
Pair 1	AWQI – PWQI	0.2360	3.0444	0.5558	-0.9008	1.3728	0.425	29	0.674
MODEL 2									
Pair 1	AWQI – PWQI1	-0.8042	3.2724	0.6544	-2.1549	0.5466	-1.229	24	0.231
MODEL 3									
Pair 1	AWQI – PWQI1	-0.0380	2.8731	0.6424	-1.3827	1.3065	-0.059	19	0.953
MODEL 4									
Pair 1	AWQI – PWQI1	0.1618	3.0753	0.7940	-1.5413	1.8648	0.204	14	0.841
MODEL 5									
Pair 1	AWQI – PWQI2	-0.0422	3.5582	0.6496	-1.3709	1.2864	-0.065	29	0.949

MODEL 6									
Pair 1	AWQI – PWQI2	0.5056	4.2058	0.8412	-1.2305	2.2417	0.601	24	0.553
MODEL 7									
Pair 1	AWQI – PWQI2	0.3811	4.0213	0.8992	-1.5009	2.2632	0.424	19	0.676
MODEL 8									
Pair 1	AWQI – PWQI2	-0.2107	2.9120	0.7519	-1.8233	1.4019	-0.280	14	0.783
MODEL 9									
Pair 1	AWQI – PWQI3	-0.7752	4.3516	0.7945	-2.4001	0.8497	-0.976	29	0.337
MODEL 10									
Pair 1	AWQI – PWQI3	0.0065	5.1531	1.0306	-2.1206	2.1336	0.006	24	0.995
MODEL 11									
Pair 1	AWQI – PWQI2	0.3811	4.0213	0.8992	-1.5009	2.2632	0.424	19	0.676
MODEL 12									
Pair 1	AWQI – PWQI3	0.1615	4.0035	1.0337	-2.0556	2.3785	0.156	14	0.878

5.1.3.5. Accuracy Check of Model

As all the models are accepted in the validation process through a t-test, so to identify the most accurate model out of these 12 models, Root Mean Square Error (RMSE) calculation is done using above mentioned equation.

Table 5.15. Model wise RMSE values.

MODEL NO.	RMSE
1 (70:30)	3.0026
2 (75:25)	3.3056
3 (80:20)	2.8007
4 (85:15)	2.9754
5 (70:30)	3.4986
6 (75:25)	4.1518
7 (80:20)	3.5589
8 (85:15)	2.8211
9 (70:30)	4.3482
10 (75:25)	5.0490
11 (80:20)	4.2979
12 (85:15)	3.8711

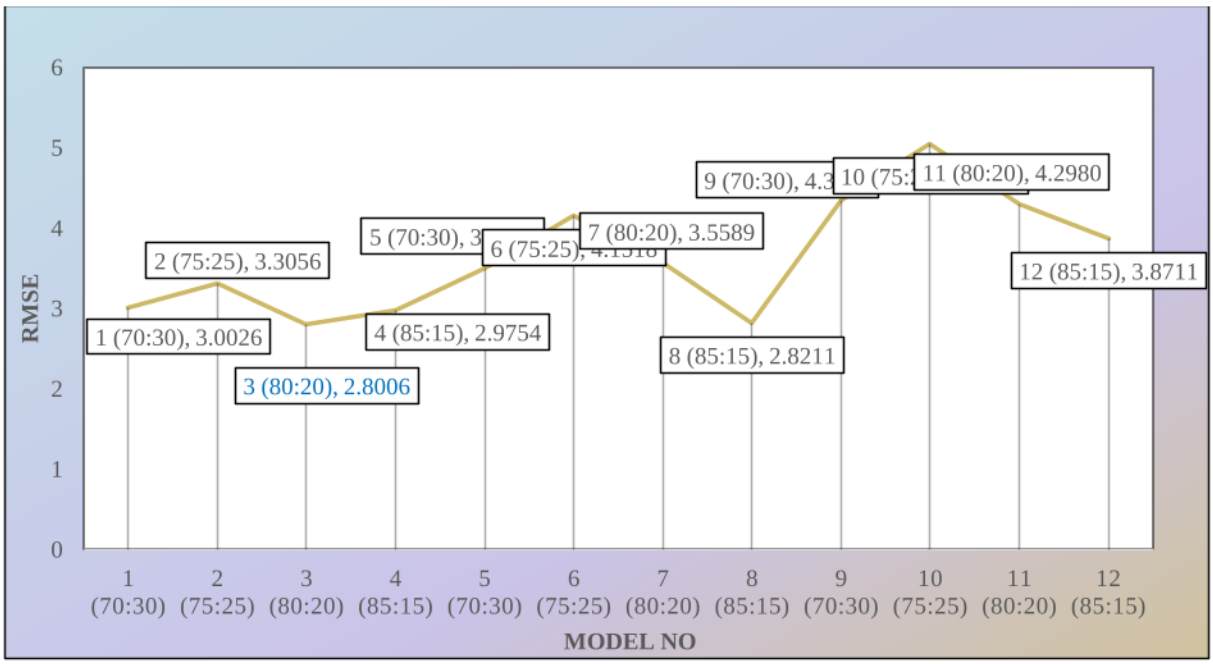


Fig. 5.3. Graphical representation of model wise RMSE value.

Out of 12 models, 12 models are capable of predicting the WQI value accurately. But among the 12 models with RMSE value 2.8006 is less as compared to the other values. To cross-check the dataset distribution ratio, out of 100 numbers samples, 60 numbers, 70 numbers, 80 numbers and 90 numbers samples are considered and the same process has been followed to prepare the MLR model. It is found that in the case of 60 and 70 numbers of samples, the dataset ratio obtained is 80:20, In the case of 80 numbers and 90 numbers of datasets, the ratios obtained are 85:15 and 75:25, respectively. Out of all these 5 sampling cases, 3 sampling cases give a dataset ratio of 80:20. Therefore, this ratio may provide better results. So the following model is proposed in this study. The input parameters are TDS, Turbidity, Mn, Fe, Cl, and TH. The dataset combination is 80:20.

5.1.4. Discussion

A hidden assumption of causality underlies the stepwise MLR approach, and the selection of independent and dependent variables in the model is essential. The independent variable is a moving force, and the dependent variable needs to be explained (Ewaid et al. 2018).

The study's goal was to create a new WQI with fewer variables but retaining a considerable amount of information, therefore it was exactly what was intended when there was no discernible difference between the original and new WQI. There are several benefits that can result from lowering the number of variables examined for the WQI, including, perhaps, financial savings. The reduction in the eclipse effect, which is defined as the attenuation of the impact of the variation of one parameter due to the stability of the others, is another illustration of the benefit brought about by the reduction in the number of independent variables in the WQI calculation.

Using this multiple linear regression (MLR) methodology, (Ewaid et al., 2018) also obtained results from a new WQI that can be compared to the original WQI in his study about the Tigre River in Baghdad. This shows that the new equation modeled by them can be used to monitor and predict the quality of the river's waters. A comparable methodology was employed in other research, including those by (Eregno et al. 2013; Valentini et al. 2021; Chenini and Khemiri, 2009), to evaluate the quality of surface and subsurface water. These investigations witness the significance of the approach and support its practicality.

CHAPTER – VI

6.1. Conclusions

In this research, the water quality has been analyzed by objective weight WQI approach like Entropy Weight WQI and CRITIC Weight WQI and also for predicting the groundwater quality by using MLR approach. The following conclusions have been obtained from this study:

- In the Entropy weight WQI approach, almost 50% of samples have been classified as Poor, 4% of samples have been classified as very poor, 41% of samples have been classified as Good for drinking and only 5% of samples are of excellent quality for drinking. No samples which are unsuitable for drinking have been found in this study.
  Similarly, in the CRITIC Weight WQI approach water quality includes excellent (8 % samples), good (47% samples), poor (44% samples) and very poor (1% samples) from 100 samples. In this method also no unsuitable samples have been found.
- Comparison between Entropy and CRITIC weight WQI has been done on the basis of ambiguity and eclipsing. From this study, it is obtained that the Entropy weight approach based on the dataset has given a better outcome than the CRITIC weight approach.
- As the dataset is not normally distributed, Spearman correlation has been used and it is found that TDS and chloride are highly correlated with WQI in this study area and pH is a less correlated parameter with WQI.
- To prepare the model, four dataset distribution ratios have been used i.e. 70:30, 75:25, 80:20 and 85:15 and three input parameters combination out of which one combination is TDS, Turbidity, Fe, Mn, Total Hardness and Cl with WQI value, second combination is pH, TDS, Turbidity, Fe, Mn, and Total Hardness with WQI value and the third one is TDS, Turbidity, Fe, Mn, and Total Hardness with WQI value. Considering these combinations, 12 models have been found. Out of all these models, Model 3 (distribution ratio 80:20) has been proposed which shows the lowest RMSE value.
- Less number of parameters will provide financial benefits also.
- However, according to the dataset, after analyzing all these things and concluded that the 80:20 dataset distribution ratio may provide a better result. But for better representation more no of dataset and more no of parameters will be required in future for further study.
- This type of model will help to predict the groundwater quality in this area.

6.2. Limitations

- A limited number of datasets and a limited number of parameters do not provide a clear representation of water quality for any particular area always. More no of parameters and more no of samples may provide a better water quality presentation.

- In this study, only two WQI methods have been applied. As several methods are available for WQI analysis so in future if more number of water quality index methods is applied then it will be better for the conclusive idea about water quality status related to this area.

- Due to the introduction of an intelligent approach, if machine learning approach modeling is introduced in future, then it may provide better modeling for water quality assessment related to this area

References

Adimalla, N. (2021). Application of the Entropy Weighted Water Quality Index (EWQI) and the Pollution Index of Groundwater (PIG) to Assess Groundwater Quality for Drinking Purposes: A Case Study in a Rural Area of Telangana State, India. *Archives of Environmental Contamination and Toxicology*, Vol. 80, Issue 1, pp. 31–40. https://doi.org/10.1007/s00244-020-00800-4

Aouiti, S., Azaza, F. H., Fetheddine, and, Melki, E., Hamdi, M., Celico, F., and Zammouri, M. (2021). Groundwater quality assessment for different uses using various water quality indices in semi-arid region of central Tunisia. *Environmental Science and Pollution Research.* Vol. 28, pp - 46669–46691. https://doi.org/10.1007/s11356-020-11149-5/Published

Banda, T. D., and Kumarasamy, M. (2020). Application of multivariate statistical analysis in the development of a surrogate water quality index (WQI) for South African watersheds. *Water*, Vol. 12, Issue 6, Art. no. 1584. https://doi.org/10.3390/W12061584

Drinking Water Specification, IS 10500 (2012), Bureau of Indian Standard, New Delhi, India.

Chen, T., Zhang, H., Sun, C., Li, H., and Gao, Y. (2018). Multivariate statistical approaches to identify the major factors governing groundwater quality. *Applied Water Science*, Vol. 8, Issue 7, pp. 1–6. https://doi.org/10.1007/s13201-018-0837-0

Chenini, I., and Khemiri, S. (2009). Evaluation of ground water quality using multiple linear regression and structural equation modeling. *International Journal of Environmental Science and Technology*, Vol. 6, Issue 3, pp. 509–519. https://doi.org/10.1007/BF03326090

Chidiac, S., El Najjar, P., Ouaini, N., El Rayess, Y., and El Azzi, D. (2023). A comprehensive review of water quality indices (WQIs): history, models, attempts and perspectives. In *Reviews in Environmental Science and Biotechnology*, Vol. 22, Issue 2, pp. 349-395. https://doi.org/10.1007/s11157-023-09650-7

Das, C. R., and Das, S. (2023). Acceptability of MEREC criteria compared to existing weighted WQI models to assess coastal groundwater quality in eastern India. *Journal of Coastal Conservation*, Vol. 27, Issue 5, Art. no. 44. https://doi.org/10.1007/s11852-023-00975-7

Das, C. R., and Das, S. (2024). Coastal groundwater quality prediction using objective-weighted WQI and machine learning approach. *Environmental Science and Pollution Research*, Vol. 31, pp. 19439-19457. https://doi.org/10.1007/s11356-024-32415-w

Das, C. R., Das, S., and Panda, S. (2023). MLR index–based principal component analysis to investigate and monitor probable sources of groundwater pollution and quality in coastal areas: a case study in East India. *Environmental Monitoring and Assessment*, Vol. 195, Issue 10, Art. no. 1158. https://doi.org/10.1007/s10661-023-11804-7

Das, D. (2018). Hydrogeochemical assessment of Groundwater in Sagar Island Region, South 24-Parganas, West Bengal, India. *International Journal of Advanced Scientific Research and Management*, Vol. 3, Issue 11. pp. 99-115.

District Environment Plan (2023), South 24 Parganas District, West Bengal, India, Office of the District Magistrate ans Collector, South 24 Parganas, Alipore, Kolkata, West Bengal, India. http://efaidnbmnnnibpcajpcglclefindmkaj/https://s24pgs.gov.in/assets/webdoc/mediaupload/9b4cf0bf199e0092a53592230cf7fd37.pdf

District Survey Report (2018), South 24 Parganas District, West Bengal, India, West Bengal Mineral Development And Trading Corporation Limited (A Govt. of West Bengal Undertaking). http://efaidnbmnnnibpcajpcglclefindmkaj/https://s24pgs.gov.in/assets/webdoc/mediaupload/x24d2d7b4a520fc497650fc11971cfa5e.pdf

Egbueri, J. C., Ameh, P. D., and Unigwe, C. O. (2020). Integrating entropy-weighted water quality index and multiple pollution indices towards a better understanding of drinking water quality in Ojoto area, SE Nigeria. *Scientific African*, Vol. 10, art. no. 00644. https://doi.org/10.1016/j.sciaf.2020.e00644

Elbeltagi, A., Deng, J., Wang, K., Malik, A., and Maroufpoor, S. (2020). Modeling long-term dynamics of crop evapotranspiration using deep learning in a semi-arid environment. *Agricultural Water Management*, Vol. 241, Art. No.106334. https://doi.org/10.1016/j.agwat.2020.106334

Eregno, F. E., Nilsen, V., Grøndahl-Rosado, R., Seidu, R., Myrmel, M., and Heistad, A. (2013). Multiple Linear Regression Models for Estimating Microbial Load in the Drinking Water Source Case from Glomma River , M.S Thesis, Department of Mathematical Sciences and Technology, Norwegian University of Life Sciences, Norway.

Ewaid, S. H., Abed, S. A., and Kadhum, S. A. (2018). Predicting the Tigris River water quality within Baghdad, Iraq by using water quality index and regression analysis. *Environmental Technology and Innovation*, Vol. 11, pp. 390–398. https://doi.org/10.1016/j.eti.2018.06.013

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, Vol. 11, Issue 1, pp. 127–138. https://doi.org/10.1007/s13762-012-0149-0

Gibrilla, A., Bam, E. K. P., Adomako, D., Ganyaglo, S., Osae, S., Akiti, T. T., Kebede, S., Achoribo, E., Ahialey, E., Ayanu, G., and Agyeman, E. K. (2011). Application of Water Quality Index (WQI) and Multivariate Analysis for Groundwater Quality Assessment of the Birimian and Cape Coast Granitoid Complex: Densu River Basin of Ghana. *Water Quality, Exposure and Health*, Vol. 3, Issue 2, pp. 63–78. https://doi.org/10.1007/s12403-011-0044-9

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*, Vol. 76, Isssue 7, Art. no. 269. https://doi.org/10.1007/s12665-017-6589-6

Govindaraju, P., Balasubrahmanian, A., Nagaraju, D., and Krishnamurthy, V. (2022). Water Quality Index (WQI) to Evaluate Groundwater Quality in Chickmagaluru District, South Karnataka, India. *Current World Environment*, Vol. 17, Issue 1, pp. 197–212. https://doi.org/10.12944/cwe.17.1.18

Jang, W. S., Engel, B., and Yeum, C. M. (2020). Integrated environmental modeling for efficient aquifer vulnerability assessment using machine learning. *Environmental Modelling and Software*, Vol. 124, Art. no. 104602. https://doi.org/10.1016/j.envsoft.2019.104602

Jaydhar, A. K., Chandra Pal, S., Saha, A., Islam, A. R. M. T., and Ruidas, D. (2022). Hydrogeochemical evaluation and corresponding health risk from elevated arsenic and fluoride contamination in recurrent coastal multi-aquifers of eastern India. *Journal of Cleaner Production*, Vol. 369, Art.no. 133150. https://doi.org/10.1016/j.jclepro.2022.133150

Jaydhar, A. K., Pal, S. C., Saha, A., Islam, A. R. M. T., Ruidas, D., and Costache, R. (2024). Hydrogeochemical properties of groundwater and associated human health hazards in coastal multiaquifers of India. *Environmental Science and Pollution Research*, Vol. 31, Issue 12, pp. 18054–18073. https://doi.org/10.1007/s11356-023-27765-w

Jesuraja, K., Selvam, S., and Murugan, R. (2021). GIS-based assessment of groundwater quality index (DWQI and AWQI) in Tiruchendur Coastal City, Southern Tamil Nadu, India. *Environmental Earth Sciences*, Vol. 80, Issue 7, pp. 1–17. https://doi.org/10.1007/s12665-021-09542-5

Jianhua, W., Peiyue, L., and Hui, Q. (2011). Groundwater quality in jingyuan county, a semi-humid area in northwest China. *E-Journal of Chemistry*, Vol. 8, Issue 2, pp. 787–793. https://doi.org/10.1155/2011/163695

Kumar, P. G. D., Viswanath, N. C., Cyrus, S., and Abraham, B. M. (2020). Mixing data for multivariate statistical study of groundwater quality. *Environmental Monitoring and Assessment*, Vol. 192, Issue 8, Art. No. 506. https://doi.org/10.1007/s10661-020-08465-1

Mahapatra, S. S., Sahu, M., Patel, R. K., and Panda, B. N. (2012). Prediction of Water Quality Using Principal Component Analysis. *Water Quality, Exposure and Health*, Vol. 4, Issue 2, pp. 93–104. https://doi.org/10.1007/s12403-012-0068-9

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

Ram, A., Tiwari, S. K., Pandey, H. K., Chaurasia, A. K., Singh, S., and Singh, Y. V. (2021). Groundwater quality assessment using water quality index (WQI) under GIS framework. *Applied Water Science*, Vol. 11, Issue 2, pp. 1–20. https://doi.org/10.1007/s13201-021-01376-7

Sajib, A. M., Diganta, M. T. M., Rahman, A., Dabrowski, T., Olbert, A. I., and Uddin, M. G. (2023). Developing a novel tool for assessing the groundwater incorporating water quality index and machine learning approach. *Groundwater for Sustainable Development*, Vol. 23, Art. no. 101049. https://doi.org/10.1016/j.gsd.2023.101049

Satish Chandra, D., Asadi, S. S., and Raju, M. V. S. (2017). Estimation of water quality index by weighted arithmetic water quality index method: A model study. *International Journal of Civil Engineering and Technology*, Vol. 8, Issue 4, pp. 1215–1222.

Sutadian, A. D., Muttil, N., Yilmaz, A. G., and Perera, B. J. C. (2016). Development of river water quality indices—a review. *Environmental Monitoring and Assessment*, Vol. 188, Issue 1, pp. 1–29. https://doi.org/10.1007/s10661-015-5050-0

Todd, D. K., and Mays, L. W. (2005). Groundwater hydrology. John Wiley and Sons.

Uddin, M. G., Nash, S., Rahman, A., and Olbert, A. I. (2022). A comprehensive method for improvement of water quality index (WQI) models for coastal water quality assessment. *Water Research*, Vol. 219, Art. no. 118532. https://doi.org/10.1016/j.watres.2022.118532

Valentini, M., dos Santos, G. B., and Muller Vieira, B. (2021). Multiple linear regression analysis (MLR) applied for modeling a new WQI equation for monitoring the water quality of Mirim Lagoon, in the state of Rio Grande do Sul—Brazil. *SN Applied Sciences*, Vol. 3, Issue 1, Art. no. 70. https://doi.org/10.1007/s42452-020-04005-1

Wang, D., Wu, J., Wang, Y., and Ji, Y. (2020). Finding High-Quality Groundwater Resources to Reduce the Hydatidosis Incidence in the Shiqu County of Sichuan Province, China: Analysis, Assessment, and Management. *Exposure and Health*, Vol. 12, Issue 2, pp. 307–322. https://doi.org/10.1007/s12403-019-00314-y

Wu, J., Li, P., Wang, D., Ren, X., and Wei, M. (2020). Statistical and multivariate statistical techniques to trace the sources and affecting factors of groundwater pollution in a rapidly growing city on the Chinese Loess Plateau. *Human and Ecological Risk Assessment*, Vol. 26, Issue 6, pp. 1603–1621. https://doi.org/10.1080/10807039.2019.1594156

Zhang, Q., Qian, H., Xu, P., Hou, K., and Yang, F. (2021). Groundwater quality assessment using a new integrated-weight water quality index (IWQI) and driver analysis in the Jiaokou Irrigation District, China. *Ecotoxicology and Environmental Safety*, Vol. 212, Art. no. 111992. https://doi.org/10.1016/j.ecoenv.2021.111992

Zhang, Q., Xu, P., and Qian, H. (2020). Groundwater Quality Assessment Using Improved Water Quality Index (WQI) and Human Health Risk (HHR) Evaluation in a Semi-arid Region of Northwest China. *Exposure and Health*, Vol. 12, Issue 3, pp. 487–500. https://doi.org/10.1007/s12403-020-00345-w