A PROJECT REPORT

ON

"COMPARATIVE ASSESSMENT OF WATER QUALITY USING MULTIVARIATE STATISTICAL ANALYSIS: A CASE STUDY OF TIRAP OCP, MAKUM COALFIELD, TINSUKIA, ASSAM"

Submitted in partial fulfilment of the requirements

for the award of the degree of

MASTER OF TECHNOLOGY

In

WATER RESOURCE ENGINEERING

Under

ASSAM SCIENCE AND TECHNOLOGY UNIVERSITY

2023 – 2025

Submitted by:

Name – Kritika Gayan

Roll No – PG/C/23/31

ASTU Roll No – 230620061012

MTECH 3rd SEMESTER

Under the guidance of

DR. TRIPTIMONI BORAH

Professor

DEPARTMENT OF CIVIL ENGINEERING

ASSAM ENGINEERING COLLEGE

JALUKBARI, GUWAHATI - 781013

ASSAM ENGINEERING COLLEGE

JALUKBARI, GUWAHATI – 781013

CERTIFICATE OF SUPERVISOR

This is to certify that the work contained in the report entitled **"Comparative Assessment of Water Quality using Multivariate Statistical Analysis: A Case Study of Tirap OCP, Makum Coalfield, Tinsukia, Assam"** submitted by-

KRITIKA GAYAN (PG/C/23/31)

A Student of M-Tech 3RD Semester to the Department of Civil Engineering, Assam Engineering College for the successful completion of the course **CEW202321- MINI PROJECT** as a partial fulfilment of the degree in **Master of Technology in Water Resource Engineering**, Civil Engineering, has been carried out under my guidance and supervision.

Dr. Triptimoni Borah

Date:

Professor,
Department Of Civil Engineering
Assam Engineering College
Jalukbari,781013- India

ASSAM ENGINEERING COLLEGE

JALUKBARI, GUWAHATI – 781013

CERTIFICATE OF HEAD OF THE DEPARTMENT

This is to certify that the work contained in the report entitled **“Comparative Assessment of Water Quality using Multivariate Statistical Analysis: A Case Study of Tirap OCP, Makum Coalfield, Tinsukia, Assam”** submitted by-

KRITIKA GAYAN (PG/C/23/31)

A Student of M-Tech 3RD Semester to the Department of Civil Engineering, Assam Engineering College for the successful completion of the course **CEW202321- MINI PROJECT** as a partial fulfillment of the degree in **Master of Technology in Water Resource Engineering**, Civil Engineering, has been carried out as per university protocols. The dissertation which is based on candidate’s own work has not been submitted to any other institution in any form

Dr. Jayanta Pathak

Date:

Professor & Head

Department Of Civil Engineering
Assam Engineering College
Jalukbari,781013 - India

DECLARATION BY THE CANDIDATE

I, a student of the Department of Water Resource Engineering, Civil Engineering, Assam Engineering College, hereby declare that we have compiled this report on the topic titled **"Comparative Assessment of Water Quality using Multivariate Statistical Analysis: A Case Study of Tirap OCP, Makum Coalfield, Tinsukia, Assam"** in 3rd Semester as a part of my M. Tech curriculum.

I also declare that the same report or any substantial portion of this report has not been submitted anywhere else as part of any requirements for any degree/diploma etc.

Date:

NAME	ROLL NO.	SIGNATURE
Kritika Gayan	230620061012	

ACKNOWLEDGEMENT

It gives me a great sense of pleasure to present the report on **“Comparative Assessment of Water Quality using Multivariate Statistical Analysis: A Case Study of Tirap OCP, Makum Coalfield, Tinsukia, Assam”** completed during my **M. Tech 3rd Semester**.

I would like to express my gratitude and would like to acknowledge **Dr Triptimoni Borah**, Professor, Civil Engineering Department of Assam Engineering College, Jalukbari, Guwahati for her able guidance and constant encouragement in every step of the project. Her, sincerity, thoroughness and perseverance have been a constant source of inspiration for us. It is only her cognizant efforts that my endeavors have seen light of the day.

I would also want to take this opportunity to thank the entire staff and the faculty members of Civil Engineering Department, Assam Engineering College for their support and cooperation during the development of my project.

Finally, I recognize the participation of my friends to my project completion. Also, this project would not have been possible without the assistance of my seniors, whose benignant guidance helped me throughout the project. It was a great source of motivation and helped me keep the momentum alive throughout this study.

ABSTRACT

This study evaluates the impact of mining activities on the water quality of surface, effluent, and drinking water sources in and around the Tirap Open Cast Project (OCP) in the Makum Coalfield, Tinsukia District, Assam. Multivariate Statistical Analysis (MSA) was employed, with a focus on **Hierarchical Cluster Analysis (HCA)**, to classify water samples based on physicochemical and microbial parameters collected over three months (October, November, and December) for the year 2021. The analysis identified significant spatial and temporal variations in water quality, with clusters revealing pollution hotspots linked to mining activities.

Results indicate elevated contamination levels in specific sites due to acid mine drainage, heavy metal leaching, and sedimentation. These findings highlight the effectiveness of HCA in identifying pollution sources and guiding water management strategies. However, temporal scope and parameter coverage limitations suggest the need for expanded research. The study provides actionable insights for sustainable water resource management and emphasizes the importance of mitigating the environmental impacts of coal mining activities.

Keywords: Tirap OCP, Water quality parameters, Hierarchical Cluster Analysis (HCA), Surface water, Effluent water, Drinking water.

CONTENTS

CHAPTER TITLE	PAGE NO.
1 INTRODUCTION	1 - 3
1.1 OVERVIEW	
1.1.1 Coal mining impacts on water quality	
1.2 OBJECTIVES	
1.3 LIMITATIONS	
2 LITERATURE REVIEW	4 - 7
3 METHODOLOGY	8 - 22
3.1 STUDY AREA	
3.2 DATA COLLECTION	
3.3 DATA ANALYSIS	
3.3.1 Multivariate statistical analysis	
3.3.2 Hierarchical statistical analysis	
4 RESULTS AND DISCUSSIONS	23 - 43
4.1 HIERARCHICAL CLUSTER ANALYSIS FOR SURFACE WATER	
4.1.1 Results showing surface water analysis using HCA for October month	
4.1.2 Results showing surface water analysis using HCA for November month	
4.1.3 Results showing surface water analysis using HCA for December month	
4.2 HIERARCHICAL CLUSTER ANALYSIS FOR EFFLUENT WATER	
4.2.1 Results showing effluent water analysis using HCA for October month	
4.2.2 Results showing effluent water analysis using HCA for November month	
4.2.3 Results showing effluent water analysis using HCA for December month	
4.3 HIERARCHICAL CLUSTER ANALYSIS FOR EFFLUENT WATER	

4.3.1 Results showing effluent water analysis using HCA for October month 4.3.2 Results showing effluent water analysis using HCA for November month 4.3.3 Results showing effluent water analysis using HCA for December month 4.4 DISCUSSIONS 4.4.1 Surface water 4.4.2 Effluent water 4.4.3 Drinking water	
5 CONCLUSIONS 5.1 OVERVIEW 5.2 OUTCOMES 5.3 SCOPE FOR FUTURE RESEARCH	44 - 45
6 REFERENCES	46

LIST OF TABLES

TABLE TITLE	PAGE NO.
3.1 Parameters for Surface water samples	11
3.2 Parameters for Effluent water samples	14
3.3 Parameters for Drinking water samples	17
4.1 Processing Summary (Surface water in October month)	23
4.2 Similarity matrix for surface water in October month	23
4.3 Ward's Linkage for Surface Water in October month	23
4.4 Processing Summary (Surface water in November month)	25
4.5 Similarity Matrix for surface water in November month	25
4.6 Ward's Linkage for Surface Water in November month	25
4.7 Processing Summary (Surface water in December month)	27
4.8 Similarity Matrix for surface water in December month	27
4.9 Ward's Linkage for Surface Water in December month	27
4.10 Processing Summary (Effluent water in October month)	29
4.11 Similarity Matrix for Effluent water in October month	29
4.12 Ward's Linkage for Effluent Water in October month	29
4.13 Processing Summary (Effluent water in November month)	31
4.14 Similarity Matrix for Effluent water in November month	31
4.15 Ward's Linkage for Effluent Water in November month	31
4.16 Processing Summary (Effluent water in December month)	33
4.17 Similarity Matrix for Effluent water in December month	33
4.18 Ward's Linkage for Effluent Water in December month	33
4.19 Processing Summary (Drinking water in October month)	35
4.20 Similarity Matrix for Drinking water in October month	35
4.21 Ward's Linkage for Drinking Water in October month	35
4.22 Processing Summary (Drinking water in November month)	37
4.23 Similarity Matrix for Drinking water in November month	37
4.24 Ward's Linkage for Drinking Water in November month	37
4.25 Processing Summary (Drinking water in December month)	39
4.26 Similarity Matrix for Drinking water in December month	39
4.27 Ward's Linkage for Drinking Water in December month	39

LIST OF FIGURES

FIGURE TITLE	PAGE NO.
3.1 STUDY AREA MAP, TIRAP OCP, NEC, MAKUM COALFIELDS	9
3.2 LOCATIONS OF SURFACE WATER SAMPLES	13
3.3 LOCATIONS OF EFFLUENT WATER SAMPLES	16
3.4 LOCATIONS OF DRINKING WATER SAMPLES	20
4.1 ICICLE PLOTS FOR SURFACE WATER IN OCTOBER MONTH	24
4.2 DENDROGRAM FOR SURFACE WATER SAMPLES IN OCTOBER MONTH	24
4.3 ICICLE PLOTS FOR SURFACE WATER IN NOVEMBER MONTH	26
4.4 DENDROGRAM FOR SURFACE WATER SAMPLES IN NOVEMBER MONTH	26
4.5 ICICLE PLOTS FOR SURFACE WATER IN DECEMBER MONTH	28
4.6 DENDROGRAM FOR SURFACE WATER SAMPLES IN DECEMBER MONTH	28
4.7 ICICLE PLOTS FOR EFFLUENT WATER IN OCTOBER MONTH	30
4.8 DENDROGRAM FOR EFFLUENT WATER SAMPLES IN OCTOBER MONTH	30
4.9 ICICLE PLOTS FOR EFFLUENT WATER IN NOVEMBER MONTH	32
4.10 DENDROGRAM FOR EFFLUENT WATER SAMPLES IN NOVEMBER MONTH	32
4.11 ICICLE PLOTS FOR EFFLUENT WATER IN DECEMBER MONTH	34
4.12 DENDROGRAM FOR EFFLUENT WATER SAMPLES IN DECEMBER MONTH	34
4.13 ICICLE PLOTS FOR DRINKING WATER IN OCTOBER MONTH	36
4.14 DENDROGRAM FOR DRINKING WATER SAMPLES IN OCTOBER MONTH	36
4.15 ICICLE PLOTS FOR DRINKING WATER IN NOVEMBER MONTH	38
4.16 DENDROGRAM FOR DRINKING WATER SAMPLES IN NOVEMBER MONTH	38
4.17 ICICLE PLOTS FOR DRINKING WATER IN DECEMBER MONTH	40
4.18 DENDROGRAM FOR DRINKING WATER SAMPLES IN DECEMBER MONTH	40

1 INTRODUCTION

1.1 OVERVIEW

Water quality plays a crucial role in sustaining ecosystems, supporting human health, and enabling industrial and agricultural activities. However, industrial processes, especially mining, pose significant threats to the quality of surface water, effluent water, and drinking water. Coal mining, in particular, disrupts natural ecosystems and hydrological processes, leading to a variety of environmental challenges. The **Tirap Open Cast Project (OCP)**, located in the **Makum Coalfield** of **Tinsukia District, Assam**, is one such site where coal mining activities significantly impact water resources.

Mining operations at Tirap OCP, including excavation, dewatering, and waste disposal, contribute to pollution through acid mine drainage (AMD), heavy metal contamination, sedimentation, and alteration of natural flow patterns. Acid mine drainage is a prominent issue, as the oxidation of sulphide minerals during mining leads to the generation of sulfuric acid, which lowers the pH and mobilizes toxic heavy metals into nearby water bodies. Additionally, the disposal of untreated effluents and runoff from mining areas results in elevated levels of suspended solids, salinity, and nutrient loading, degrading the water quality of the region.

To address these concerns, **Multivariate Statistical Analysis (MSA)** offers an effective framework for evaluating complex datasets and understanding spatial and temporal variations in water quality. Among the multivariate techniques, **Hierarchical Cluster Analysis (HCA)** is particularly valuable for classifying sampling sites based on similarities in water quality parameters. By identifying clusters of sites with similar characteristics, HCA provides critical insights into pollution sources and levels of contamination. This study leverages HCA to assess the water quality of surface, effluent, and drinking water in and around the Tirap OCP, considering physicochemical and microbial parameters over different months.

1.1.1 Coal Mining Impacts on Water Quality

Coal mining significantly impacts water quality through the following mechanisms:

1.1.1.1 *Acid Mine Drainage (AMD):*

The exposure of sulphide minerals to air and water during mining generates sulfuric acid, lowering the pH and increasing the solubility of toxic metals like iron, manganese, and arsenic. This process leads to the contamination of nearby water bodies.

1.1.1.2 *Heavy Metal Contamination:*

Mining and associated waste disposal result in heavy metals such as lead, cadmium, and chromium leaching into surface and groundwater systems. These metals pose severe risks to aquatic ecosystems and human health.

1.1.1.3 *Suspended Solids and Sedimentation:*

Mining activities increase the concentration of suspended solids in water bodies, leading to turbidity, reduced light penetration, and sediment deposition, which affect aquatic habitats.

1.1.1.4 *Nutrient Loading:*

Runoff from mining areas often carries excess nitrates and phosphates, causing eutrophication and depletion of dissolved oxygen in downstream water bodies.

1.1.1.5 *Alteration of Natural Flow Patterns:*

Mine dewatering and waste dumping disrupt natural hydrological systems, reducing water availability and altering flow regimes.

1.1.1.6 *Salinity Increases:*

Mining wastes often contain high concentrations of dissolved salts, which leach into water bodies, raising salinity levels and rendering water unsuitable for agricultural or drinking purposes.

1.1.1.7 *Microbial Contamination:*

Improper waste management can introduce microbial pollutants into water bodies, posing health hazards to nearby communities.

1.2 OBJECTIVES

The primary objectives of this study are:

i. ***Water Quality Assessment:*** To evaluate the quality of surface water, effluent water, and drinking water in the study area based on physicochemical and microbial parameters.

ii. ***Pollution Source Identification****:* To identify the key sources of pollution associated with mining activities and related industrial operations.
iii. ***Spatial and Temporal Analysis****:* To study seasonal variations in water quality and spatial differences among the sampling locations.
iv. ***Cluster Identification****:* To classify sampling sites using **Hierarchical Cluster Analysis (HCA)** and determine clusters representing high and low pollution levels.
v. ***Policy Recommendations****:* To provide actionable insights for sustainable water resource management and pollution control strategies in the region.

1.3 LIMITATIONS

While this study provides valuable insights into water quality in the Tirap OCP region, it has several limitations:

i. ***Temporal Scope****:* The analysis is limited to three months (October, November, December), which may not capture the full range of seasonal variations.
ii. ***Parameter Selection****:* The study focuses on selected physicochemical and microbial parameters, potentially overlooking other pollutants like hydrocarbons or emerging contaminants.
iii. ***Geographical Coverage****:* Sampling is confined to areas within a defined radius of the Tirap OCP, which may not reflect broader regional impacts.
iv. ***Data Standardization****:* Variations in sampling methods and environmental conditions might influence the accuracy and comparability of results.
v. ***External Influences****:* Agricultural runoff, industrial discharges, and other external factors beyond mining activities may confound the interpretation of results.

Addressing these limitations in future research can help achieve a more comprehensive understanding of water quality dynamics in coal mining regions. This study's findings provide a critical foundation for developing effective water management strategies and mitigating the adverse environmental effects of coal mining.

2 LITERATURE REVIEW

Tiri and Lahbari (2015) The study aimed to assess the quality of surface water in the Koudiat Medouar Watershed, East Algeria, and identify the spatial and temporal variations in water quality. It also sought to determine the sources of hydro chemical elements influencing water quality. 42 surface water samples were collected from three stations (Oued Reboa, Oued Timgad, and Basin Dam) between June 2010 and February 2011. Hydro chemical analysis to measure parameters like pH, electrical conductivity (EC), major ions (Mg, Ca, Na, K, Cl, SO_4, HCO_3, and NO_3), and temperature. Statistical methods such as **Hierarchical Cluster Analysis (HCA)** and **Analysis of Variance (ANOVA)** were employed to evaluate variations and group samples. Digital portable water analyser used for onsite measurements. Flame photometer, UV-Visible spectrophotometer, and titration methods used for laboratory analysis. Statistical analyses were performed using **STATISTICA®** software. The water in the watershed exhibited alkaline characteristics, with significant variations in hydro chemical parameters across stations and time. Two primary hydro chemical facies were identified:

- Mg-HCO_3 at Oued Reboa and Oued Timgad.
- Mg-SO_4 at the Basin Dam.

Anthropogenic activities (e.g., agriculture) and natural processes (e.g., water-rock interactions) were the main sources of hydro chemical variations. The study provides crucial insights into water quality management for the region.

Dawood (2017) The study aimed to evaluate the spatial and temporal variations in surface water quality of the Shatt Al-Arab River (Iraq) using **multivariate statistical methods**. The goal was to interpret complex water quality datasets, identify pollution sources, and propose optimal monitoring strategies for effective water resource management. The research focused on the Shatt Al-Arab River, which forms from the confluence of the Tigris and Euphrates Rivers and flows through Basrah, Iraq. Monthly water samples were collected over a four-year period (2011–2014) from five monitoring sites. Seven water quality parameters were analysed: **Dissolved Oxygen (DO)**, **Phosphate (PO4)**, **Calcium (Ca)**, **Magnesium (Mg)**, **Nitrate (NO3)**, **Chloride (Cl)**, and **Sulphate (SO4)**. **Principal Component Analysis (PCA)** used to reduce data dimensionality and identify key factors contributing to water quality variation. **Cluster Analysis (CA)** employed to classify sampling sites into clusters based on their similarity in water quality characteristics. Software like **SPSS 22** used for PCA and statistical analyses, **Minitab 16** for cluster analysis (CA). The study successfully applied PCA and CA to evaluate spatial and temporal variations in water quality. Three principal components were extracted, explaining 98.9% of the total variance. Cluster analysis grouped the five monitoring sites into two clusters:

- **Cluster 1**: Represented less polluted sites.
- **Cluster 2**: Included more polluted sites affected by domestic, industrial, and agricultural effluents.

The findings underscore the effectiveness of multivariate statistical methods in analysing complex water quality data and highlight the need to regulate wastewater discharge to protect the Shatt Al-Arab River. The results provide a foundation for future water quality monitoring and management plans.

Barrie and et. al (2023) The study aimed to evaluate the water quality of the Rokel River in Sierra Leone using **multivariate statistical approaches**. It sought to analyse the relationships between water quality parameters, identify pollution sources, and assess seasonal and spatial variations to inform water management strategies. Water samples were collected from four locations in the Rokel River during the wet (June–August) and dry (February–April) seasons. Physicochemical and microbial parameters were measured, including pH, turbidity, TDS, DO, nitrate, phosphate, fluoride, and heavy metals. **Principal Component Analysis (PCA)** used to reduce dimensionality and identify key factors influencing water quality. **Hierarchical Cluster Analysis (HCA)** applied to classify sampling sites based on parameter similarities and seasonal variations. **Water Quality Index (WQI)**: Calculated to grade water quality and identify sites requiring treatment. Kaiser-Meyer-Olkin (KMO) and Bartlett's tests were conducted to ensure data adequacy for PCA. Analysis of Variance (ANOVA) was used to test the statistical significance of seasonal variations. **SPSS 16.0** used for conducting PCA, HCA, and statistical analysis and **Microsoft Excel** used for data organization and basic computations. The study revealed significant seasonal variations in water quality, with higher pollution during the rainy season due to runoff and human activities. Parameters such as chromium, iron, phosphate, and fluoride exceeded WHO guidelines, indicating pollution from industrial and agricultural sources. The findings highlight the need for targeted interventions and improved water management strategies to ensure safe water supply and environmental sustainability.

Rautela and et. al (2023) The study aimed to evaluate the surface water quality of the Alaknanda River basin in the Indian Himalayas. It focused on seasonal variations in water quality and the influence of anthropogenic and natural factors using multivariate statistical methods and water quality indices (WQI). 44 water samples were collected across 15 sites during four seasons (monsoon, winter, spring, and summer) from 2021 to 2022. Physicochemical parameters like pH, turbidity, dissolved oxygen (DO), and total suspended solids (TSS) were measured. Water Quality Index (WQI) was calculated using established formulas. Multivariate statistical techniques, including Principal Component Analysis (PCA) and correlation matrices, identified key factors influencing water quality. Laboratory methods were applied as per APHA and BIS standards to determine the parameters. PCA and other statistical analyses were performed on standardized datasets, likely using statistical software like SPSS, MATLAB. WQI values were strongly influenced by TSS and turbidity, particularly during monsoon and summer seasons when glacier melting and intense rainfall increased sediment levels. The water quality was generally within permissible limits for drinking after sedimentation, except during high-flow periods. Anthropogenic activities (e.g., agricultural runoff, household wastewater) and natural factors (e.g., glacier melting) significantly impacted water quality. The findings provide essential data for water quality management and policy-making to improve the ecological health of the Himalayan rivers.

Kataki and Borah (2024) The study aims to evaluate seasonal and spatial variations in the water quality of Deepor Beel, a Ramsar wetland in Assam, over one year. It seeks to identify pollution sources and their effects on aquatic ecosystems and water quality using multivariate statistical techniques. Water samples were collected from nine fixed locations in Deepor Beel across four seasons (autumn, winter, spring, and summer). Physicochemical parameters (e.g., pH, dissolved oxygen, TDS, EC, BOD) and heavy metals (e.g., lead, iron) were analysed. In-situ measurements were performed using a multiparameter water quality analyser. Additional analyses were conducted in the laboratory, following WHO and BIS standards. Techniques like one-way ANOVA, Pearson correlation, and hierarchical cluster analysis (HCA) were applied to assess temporal and spatial variations in water quality. Significant seasonal variations were found in parameters like EC, TDS, BOD, and dissolved oxygen. Some parameters, such as nitrate and lead, exceeded permissible limits in certain seasons. Sampling sites were grouped into clusters based on pollution levels, with some clusters indicating higher contamination due to industrial effluents, agricultural runoff, and residential waste. A strong positive or negative correlation between various parameters highlighted anthropogenic impacts as a major source of contamination. Key pollutants, such as lead and iron, require urgent attention. The results emphasize the necessity for timely interventions to prevent further ecological degradation of the wetland.

3 METHODOLOGY

3.1 STUDY AREA

The study area spans an aerial extent of 425.68 km², encompassing a 10 km radius around the Tirap Open Cast Project (OCP) of the Makum Coalfields, located in Tinsukia District, Assam, India. The area is situated between latitudes 27°11'45" N to 27°24'05" N and longitudes 95°40'00" E to 95°54'15" E, falling within the 83M/15 topographical sheet.

This region is characterized by diverse landforms, ranging from alluvial plains to highly dissected hills and valleys. The alluvial plains, formed by the Burhi-Dihing River, the largest south-bank tributary of the Brahmaputra in Upper Assam are a prominent feature. The Burhi-Dihing River drains a basin of approximately 6,000 km², with its width varying from 300 m to 400 m in the plains. This perennial, meandering river has a sinuosity of 1.6, with a dendritic drainage pattern in its lower-order streams and a sub-parallel pattern in its higher-order streams.

The terrain to the north of NH-315 is predominantly low-relief and features flat to gently rolling landforms, dissected by small gullies that channel surface runoff into the Ledopani Nala. The Ledopani Nala, a 4th-order stream located 800 m west of the project, flows south to north and eventually joins the Burhi-Dihing River, located 4.5 km north-northwest of the project area.

Another significant watercourse, the Lekhapani Nala, also a 4th-order stream, flows in a northeasterly direction in the south and east of the project area, at a distance of about 1.6 km. Multiple tributaries of the Burhi-Dihing River, such as the Namdang Nala, Ledopani Nala, Manmau Jan, Sipijan Nala, and Tirap River (located 2.2 km north of the project area), drain the surrounding buffer zone.

The overall drainage pattern of the study area is predominantly dendritic, reflecting the natural flow dynamics of the region.

Fig 3.1: Study Area Map, Tirap OCP, NEC, Makum Coalfields

3.2 DATA COLLECTION

Water samples were collected from various locations within the study area to analyse the quality of Surface Water, Effluent Water, and Drinking Water. The collection details are as follows:

- **Surface Water**: Five samples were collected from designated surface water sources.
- **Effluent Water**: Three samples were collected from identified effluent discharge points.
- **Drinking Water**: Six samples were collected from drinking water sources.

3.2.1 Sampling:

Sampling was conducted monthly during **October**, **November**, and **December**, ensuring a comprehensive representation of seasonal variations. For all water types, a range of **physicochemical** and **microbial parameters** were analysed, including:

- pH,
- Temperature,

- Total Hardness,
- Calcium, and others.

The data for all samples were obtained from the COMPREHENSIVE REPORT ON GROUNDWATER CONDITIONS FOR TIRAP OCP, NEC, provided by **CMPDI** (Central Mine Planning and Design Institute).

The tables 3.1, 3.2 and 3.3 present the parameters and their respective values for surface water, effluent water, and drinking water during the months of October, November, and December.

Table 3.1: Parameters for Surface water samples

			October					November					December				
S. No.	Parameters	Units	Tirap River Up-Stream near Lekhapani Bridge	Tirap River Down-Stream (near Molong bangali gaon 1)	Buri Dehing River Up-Stream (Near Manmae Maichang Gaon)	Buri Dehing Down-Stream (Near Circuit House)	Mixing Zone of Buri Dehing and Ledo Pani Mallah (Near Ledo Namdang Gaon)	Tirap River Up-Stream near Lekhapani Bridge	Tirap River Down-Stream (near Molong bangali gaon 1)	Buri Dehing River Up-Stream (Near Manmae Maichang Gaon)	Buri Dehing Down-Stream (Near Circuit House)	Mixing Zone of Buri Dehing and Ledo Pani Mallah (Near Ledo Namdang Gaon)	Tirap River Up-Stream near Lekhapani Bridge	Tirap River Down-Stream (near Molong bangali gaon 1)	Buri Dehing River Up-Stream (Near Manmae Maichang Gaon)	Buri Dehing Down-Stream (Near Circuit House)	Mixing Zone of Buri Dehing and Ledo Pani Mallah (Near Ledo Namdang Gaon)
1	pH (at 25 °C)	–	7.6	7.45	6.89	6.67	6.53	7.45	6.86	7.19	7.15	6.54	7.05	6.95	7.19	7.08	7.15
2	Colour	Hazen Unit	1	1	1	1	1	1	2	2	2	5	1	1	3	4	2
3	Total Hardness as CaCO3	mg/l	44	54	32	145	360	98	110	56	68	900	105	118	68	79	715
4	Calcium as Ca	mg/l	8.81	10.42	7.21	32.06	72.14	14.4	20.8	12.8	12.8	88.2	21.5	22.5	14.5	15.6	122
5	Chloride as Cl	mg/l	4.94	4.94	4.94	9.89	14.84	4.8	15.5	6.79	6.79	13.6	5.9	16.1	7.15	7.25	15.2
6	Magnesium as Mg	mg/l	5.34	6.8	3.4	15.79	43.74	15.07	14.09	5.83	8.75	165	12.2	14.8	7.56	9.5	98.5
7	Total Dissolved Solids	mg/l	64	720	48	1069	402	112	140	91	98	1087	116	146	108	113	1045
8	Sulphate as SO4	mg/l	19.7	20.6	10	109	115	10.4	50.2	15.4	27.1	105	11.7	46.8	16.6	28.5	97
9	Fluoride	mg/l	0.2	0.35	0.2	0.2	0.25	0.2	0.2	0.2	0.2	0.47	0.2	0.2	0.2	0.2	0.47
10	Nitrate as NO3	mg/l	1.2	1.2	1.68	1.5	1.05	1.38	1.42	1.2	1.3	2.9	1.25	1.56	1.54	1.36	3.14
11	Iron as Fe	mg/l	0.02	0.02	0.02	0.02	0.02	0.21	0.33	0.16	0.33	1.33	0.28	0.38	0.19	0.37	0.97
12	Phenolic compounds as C_6H_5OH	mg/l	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
13	Zinc as Zn	mg/l	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
14	Chromium as Cr	mg/l	0.01	0.01	0.211	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

15	Copper as Cu	mg/l	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
16	Manganese as Mn	mg/l	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
17	Cadmium as Cd	mg/l	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
18	Lead as Pb	mg/l	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005
19	Selenium as Se	mg/l	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005
20	Arsenic as As	mg/l	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005
21	Mercury as Hg	mg/l	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005
22	Sodium as Na mg/l	mg/l	2	2	2	5	8	3	9	4	4	7	6	5	7	5	8
23	Potassium as K	mg/l	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1
24	Total suspended solid	mg/l	1	1	1	1	1	7	9	15	13	27	7	4	15	13	27
25	BOD,3 days @27°C as O_2	mg/l	1	1	1	1	1	2	3	4	3	4	1	1	1	1	2
26	Chemical oxygen demand as O_2	mg/l	4	4	4	4	4	16	24	36	28	36	5	4	5	6	10
27	Oil and Grease	mg/l	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
28	Temperature	°C	28.9	28.7	28.7	28.5	28.9	28.7	28.8	28.7	28.8	28.7	28.5	28.6	28.8	28.9	28.9
29	Dissolved oxygen	mg/l	6.5	6.6	6.5	6.4	6.6	6.6	6.4	6.4	6.5	6.3	6.8	6.3	6.8	6.9	6.7
30	Cyanide	mg/l	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

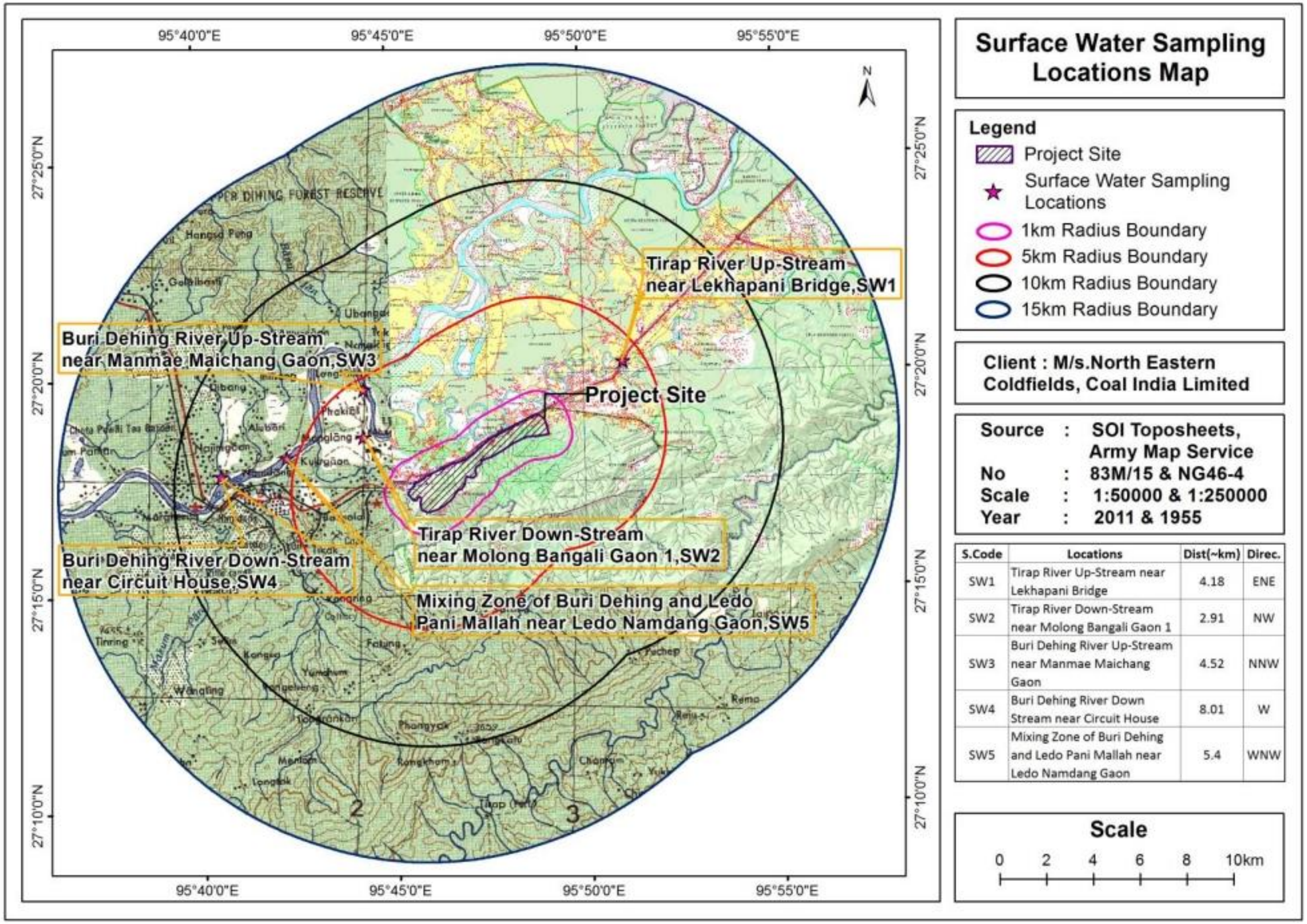


Fig 3.2: Locations of Surface water samples

Table 3.2: Parameters for Effluent water samples

S. No.	Parameters	Units	October			November			December		
			Tirap East Near Sedimentation Tank	Tirap West Near Discharge Point	Water-Sedimentation Tank of Tirap OCP	Tirap East Near Sedimentation Tank	Tirap West Near Discharge Point	Water-Sedimentation Tank of Tirap OCP	Tirap East Near Sedimentation Tank	Tirap West Near Discharge Point	Water-Sedimentation Tank of Tirap OCP
1	Colour	HU	15	5	10	5	4	1	5	5	1
2	Total suspended Solids	mg/l	122	22	45	8	26	0.1	136	52	6
3	pH @ 25°C	-	3.01	3.77	3.3	3.07	3.03	7.73	3.06	6.3	2.96
4	Dissolved Phosphate	mg/l	0.19	0.04	0.35	8.91	3.86	6.5	7.2	3.14	5.2
5	Temperature	°C	29	28.9	29	28.9	27	29	28.1	28	28.2
6	Oil & Grease	mg/l	4	4	4	4	4	4	4	4	4
7	Ammoniacal Nitrogen as N	mg/l	0.2	0.02	8.4	0.87	1.615	0.02	0.75	1.2	0.02
8	Total Kjeladahl Nitrogen as N	mg/l	0.3	0.02	12.93	1.24	2.1	0.02	1.15	1.7	0.02
9	BOD, 3 days @ 27°C as O_2	mg/l	1	10	1	1	1	1	1	1	1
10	COD as O_2	mg/l	6	49	4	4	4	4	4	4	4
11	Lead as Pb	mg/l	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
12	Chromium (Hexavalent) as Cr^{6+}	mg/l	0.105	0.1	0.6	0.1	0.1	0.415	0.1	0.1	0.21
13	Total Chromium as Cr	mg/l	0.02	0.01	0.01	0.01	0.022	0.01	0.01	0.01	0.01
14	Copper as Cu	mg/l	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
15	Zinc as Zn	mg/l	0.168	0.25	0.02	0.033	0.042	0.01	0.01	0.01	0.01
16	Boron as B	mg/l	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
17	Chloride as Cl-	mg/l	29.99	39.98	49.98	19.85	39.7	59.55	24.2	11.2	52.7
18	Fluoride as F-	mg/l	0.2	0.2	0.2	0.7	0.3	0.2	0.4	0.2	0.2
19	Sulphate as SO_4	mg/l	2173.1	988.4	1633	2137.3	1865.5	288.2	2256	58.2	1966.2
20	Sulphide as S_2-	mg/l	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
21	Arsenic as As	mg/l	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005
22	Mercury as Hg	mg/l	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
23	Cadmium as Cd	mg/l	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
24	Selenium as Se	mg/l	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005

25	Cyanide as CN	mg/l	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
26	Phenolic Compounds as C_6H_5OH	mg/l	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
27	Turbidity	NTU	70	42.7	73.6	27.2	9.4	0.1	50.7	71.9	23.3
28	Total Hardness as $CaCO_3$	mg/l	3980	1120	2140	5300	6200	0.001	2431	188	1660
29	Calcium as Ca	mg/l	336	120	160	560	360	9.4	614	42	430
30	Residual Free Chlorine	mg/l	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
31	Vanadium	mg/l	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
32	Manganese	mg/l	8.9	6.51	1.146	0.05	0.05	0.05	0.05	0.05	0.05
33	Nickel	mg/l	0.734	0.428	0.1	0.1	0.1	0.1	0.1	0.1	0.1
34	Iron	mg/l	1.9	0.34	3.05	20.49	19.17	0.02	18.2	8.25	10.6

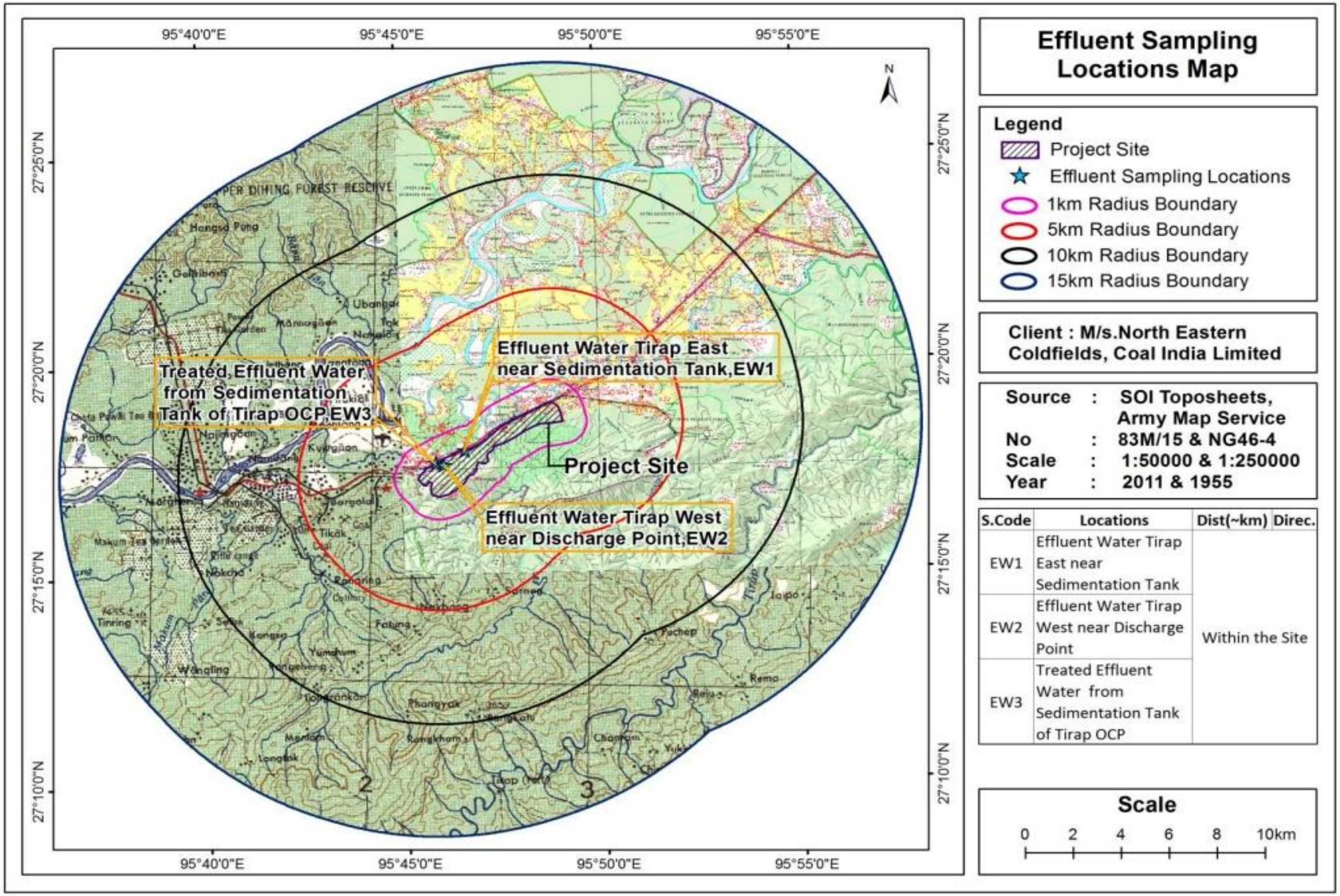


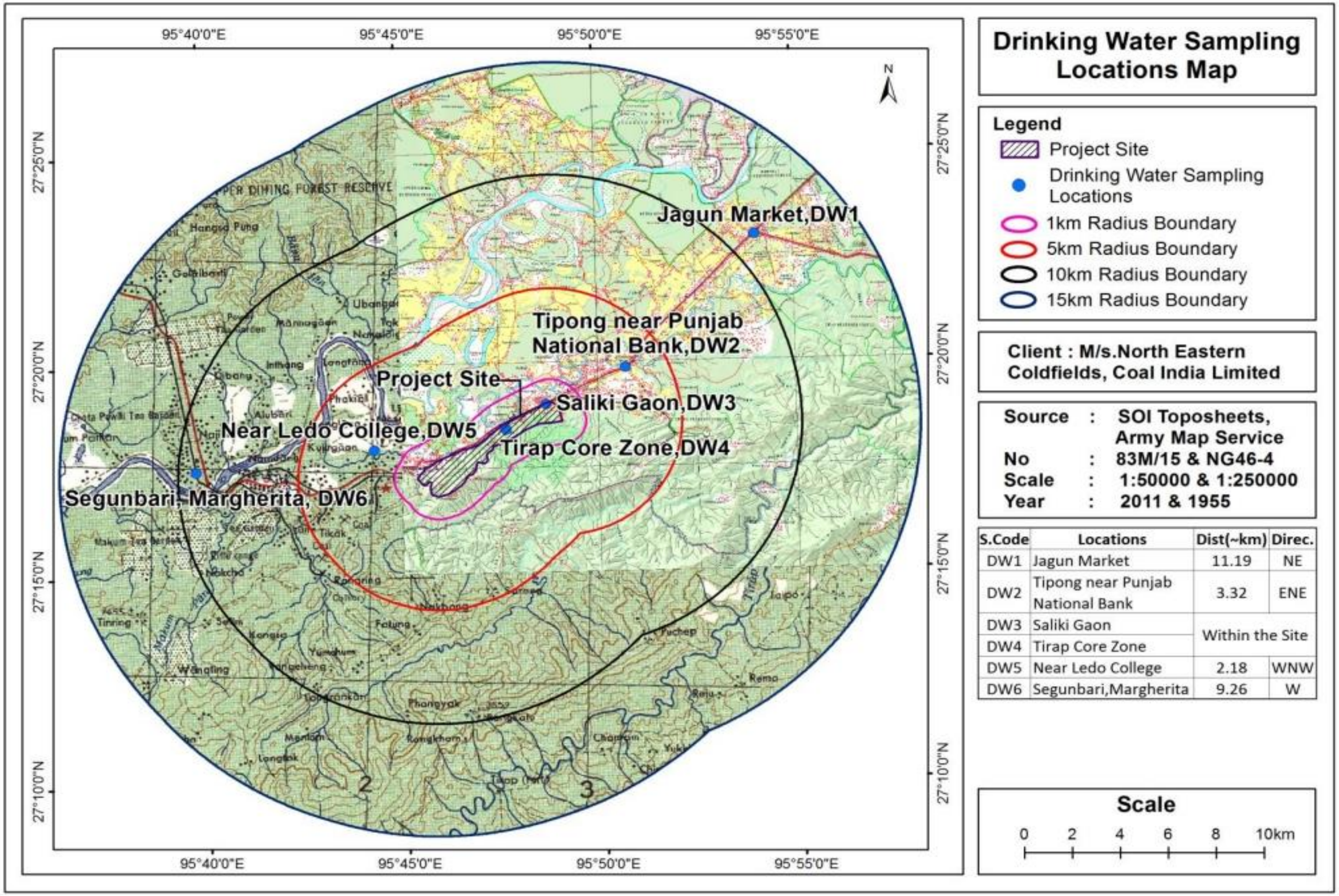
Fig 3.3: Locations of Effluent water samples

Table 3.3: Parameters for Drinking water samples

			October						November						December					
S. No.	Parameters	Units	Jagun Market	Tipong near Punjab National Bank	Saliki Gaon	Tirap Core Zone	Near Ledo College	Segunbari, Margherita	Jagun Market	Tipong near Punjab National Bank	Saliki Gaon	Tirap Core Zone	Near Ledo College	Segunbari, Margherita	Jagun Market	Tipong near Punjab National Bank	Saliki Gaon	Tirap Core Zone	Near Ledo College	Segunbari, Margherita
1	pH (at 25°C)		6.55	6.52	6.53	6.5	6.56	6.6	6.57	6.53	6.55	6.62	6.53	6.55	6.54	6.58	6.51	6.55	6.53	6.51
2	Turbidity	NTU	0.1	0.1	0.1	0.1	0.1	0.1	0.27	0.45	0.27	0.1	0.72	0.48	0.48	0.81	0.35	0.1	0.1	0.13
3	Total Hardness as $CaCO_3$	mg/l	50	48	110	32	40	66	54	48	52	46	38	102	54	40	52	40	40	72
4	Calcium as Ca	mg/l	13.62	8.81	24.04	7.21	8.81	12.82	11.2	8.82	8.82	6.41	6.41	17.6	11.22	9.62	9.62	5.61	8.01	12.82
5	Total alkalinity as $CaCO_3$	mg/l	44	66	30	30	160	34	42	66	38	36	20	54	44	76	44	54	18	58
6	Chloride as Cl	mg/l	32.66	5.93	69.28	5.93	19.79	11.87	49.5	4.85	24.3	7.76	20.4	45.6	50.47	3.96	24.74	8.9	21.77	31.67
7	Residual Free Chlorine	mg/l	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
8	Magnesium as Mg	mg/l	3.88	6.31	12.15	3.4	4.37	8.26	6.32	6.32	7.29	7.29	5.35	14.09	6.31	3.89	6.8	6.31	4.86	9.72
9	Total Dissolved Solids	mg/l	125	75	185	47	80	92	160	79	95	65	115	219	152	78	103	76	127	185
10	Sulphate as SO4	mg/l	23.3	8.8	14.6	6.9	21.6	30.7	22.8	11.2	2.49	7.17	22.7	23.6	14.72	2.4	5.43	4.32	24.2	23.95
11	Fluoride	mg/l	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2

12	Nitrate as NO3	mg/l	1.5	1.33	8.8	3.17	3.77	1.38	1.66	1.5	1.8	1.6	4.82	7.4	1.92	1.57	7.83	8.2	26.7	37.4
13	Iron as Fe	mg/l	0.02	0.02	0.02	0.02	0.02	0.02	0.68	0.25	0.75	0.02	0.16	0.11	0.31	0.56	0.85	0.02	0.02	0.14
14	Aluminium as Al	mg/l	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
15	Boron as B	mg/l	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
16	Phenolic compounds as C_6H_5OH	mg/l	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
17	Anionic Detergents as MBAS	mg/l	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
18	Zinc as Zn	mg/l	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
19	Copper as Cu	mg/l	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
20	Manganese as Mn	mg/l	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
21	Cadmium as Cd	mg/l	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
22	Lead as Pb	mg/l	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005
23	Selenium as Se	mg/l	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005
24	Arsenic as As	mg/l	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005
25	Mercury as Hg	mg/l	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005
26	Sodium as Na	mg/l	17	2	35	2	10	6	25	3	13	4	11	24	26	2	12	4	11	16
27	Potassium as K	mg/l	1	1	3	1	1	1	2	1	1	1	1	2	2	1	1	1	1	1

28	Total suspended solid	mg/l	1	1	1	1	1	1	54	97	58	1	13	9	93	321	75	1	1	27
29	Temperatur e	mg/l	28.4	28.7	28.9	29	28.8	28.5	28.8	28.7	28.8	28.8	28.7	28.8	28.7	28.8	28.7	28.6	28.8	28.6
30	Cyanide	mg/l	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
31	Chromium (Hexavalen t) as Cr^{6+}	mg/l	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1



S.Code	Locations	Dist(~km)	Direc.
DW1	Jagun Market	11.19	NE
DW2	Tipong near Punjab National Bank	3.32	ENE
DW3	Saliki Gaon	Within the Site	
DW4	Tirap Core Zone		
DW5	Near Ledo College	2.18	WNW
DW6	Segunbari,Margherita	9.26	W

Fig 3.4: Locations of Drinking water samples

3.3 DATA ANALYSIS

After collecting the data, it was organized in an Excel sheet to facilitate its insertion into SPSS. The data was then analysed using a multivariate statistical method called Hierarchical Cluster Analysis.

3.3.1 Multivariate Statistical Analysis

It is a statistical method used to analyse data that involves multiple variables simultaneously. These methods help uncover patterns, relationships, and structures in complex datasets by considering the interdependence of multiple variables.

Various common techniques for Multivariate analysis are:

1. Principal Component Analysis (PCA)
2. Hierarchical Cluster Analysis (HCA)
3. k-means Clustering
4. Discriminant Analysis
5. Factor Analysis
6. Multiple Regression Analysis
7. Canonical Correlation Analysis
8. Multivariate Analysis of Variance (MANOVA)

3.3.2 Hierarchical Statistical Analysis

Hierarchical Cluster Analysis is a multivariate statistical method used for grouping objects (e.g., data points, variables, or sampling sites) into clusters based on a distance or similarity measure (e.g., Euclidean distance, Manhattan distance, or correlation coefficients). Unlike k-means clustering, HCA does not require specifying the number of clusters in advance. The process involves organizing data into a tree-like structure called a dendrogram, which visually represents the hierarchical relationships among the objects, visually illustrates how objects are grouped at various levels of similarity.

There are various types of Linkage methods based on which clusters are merged:

1. Single Linkage (Nearest Neighbour)
   - Distance between the closest points of two clusters.

2. Complete Linkage (Farthest Neighbour)
    - Distance between the farthest points of two clusters.
3. Average Linkage
    - Average distance between all points in two clusters.
4. Centroid Linkage
    - Distance between the centroids (means) of two clusters.
5. Ward's Method
    - Minimizes the total within-cluster variance.

Advantages:

- Suitable for datasets where the number of clusters is unknown.
- Provides a visual representation of clustering through the dendrogram.
- Captures hierarchical relationships between clusters.

Limitations:

- Computationally intensive for large datasets.
- Sensitive to noise and outliers.
- Results can vary depending on the chosen linkage method and distance metric.

In this report, Hierarchical Cluster Analysis (HCA) was employed to perform Multivariate Statistical Analysis on Surface Water, Effluent Water, and Drinking Water samples collected over three months (October, November, and December).

The analysis utilized Ward's Linkage Method to measure the distances between clusters, as it is particularly effective in handling multivariate data by minimizing within-cluster variance. The Squared Euclidean Distance Method was applied to evaluate the similarity between individual data points. The analysis was conducted using the **IBM SPSS** software (free trial version), which was downloaded from its official website.

For each case (Surface Water, Effluent Water, and Drinking Water), three separate dendrograms were generated to visually represent the clustering of samples based on various physicochemical and microbial parameters analysed for the three months. These dendrograms provide insight into the relationships and similarities between sampling points, enabling the identification of meaningful clusters.

4 RESULTS AND DISCUSSIONS

4.1 Hierarchical Cluster Analysis for Surface water:

4.1.1 Results showing Surface water analysis using Hierarchical Cluster Analysis for October month

Table 4.1: Table showing Processing Summary

Case Processing Summary					
Cases					
Valid		Missing		Total	
N	Percent	N	Percent	N	Percent
5	100.0%	0	0.0%	5	100.0%
a. Squared Euclidean Distance used					
b. Ward linkage					

Table 4.2: Similarity Matrix for surface water in October month

Proximity Matrix					
Case	Squared Euclidean Distance				
	1: Tirap River Up-Stream near Lekhapani Bridge	2: Tirap River Down-Stream (near Molong bangali gaon 1)	3: Buri Dehing River Up-Stream (Near Manmae Maichang Gaon)	4: Buri Dehing Down-Stream (Near Circuit House)	5: Mixing Zone of Buri Dehing and Ledo Pani Mallah (Near Ledo Namdang Gaon)
1: Sample 1	.000	15.777	15.514	25.729	40.141
2: Sample 2	15.777	.000	24.840	29.096	40.207
3: Sample 3	15.514	24.840	.000	20.853	45.471
4: Sample 4	25.729	29.096	20.853	.000	22.372
5: Sample 5	40.141	40.207	45.471	22.372	.000

The degree of similarity (or dissimilarity) between different sampling sites using squared Euclidean distance is quantified in this table. The most dissimilar samples are Sample 3 and 5.

Table 4.3: Ward linkage for Surface water in October month

Agglomeration Schedule						
Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	1	3	7.757	0	0	2
2	1	2	18.710	1	0	4
3	4	5	29.896	0	0	4
4	1	4	56.000	2	3	0

The table highlights the step-by-step process of forming clusters, showing how individual or grouped samples combine based on Ward's method. This table consist of four stages with highest coefficient for stage 4.

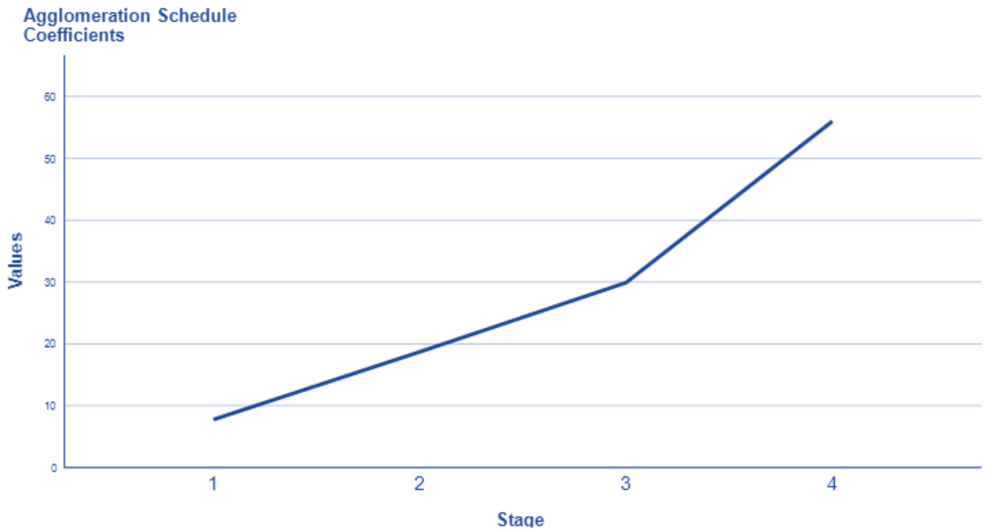


Fig 4.1: Icicle plots for Surface water in October month

The plots of agglomeration schedule coefficients (icicle plots) indicate the strength of cluster formation at each stage, with sharp increase in coefficients suggesting optimal cluster points.

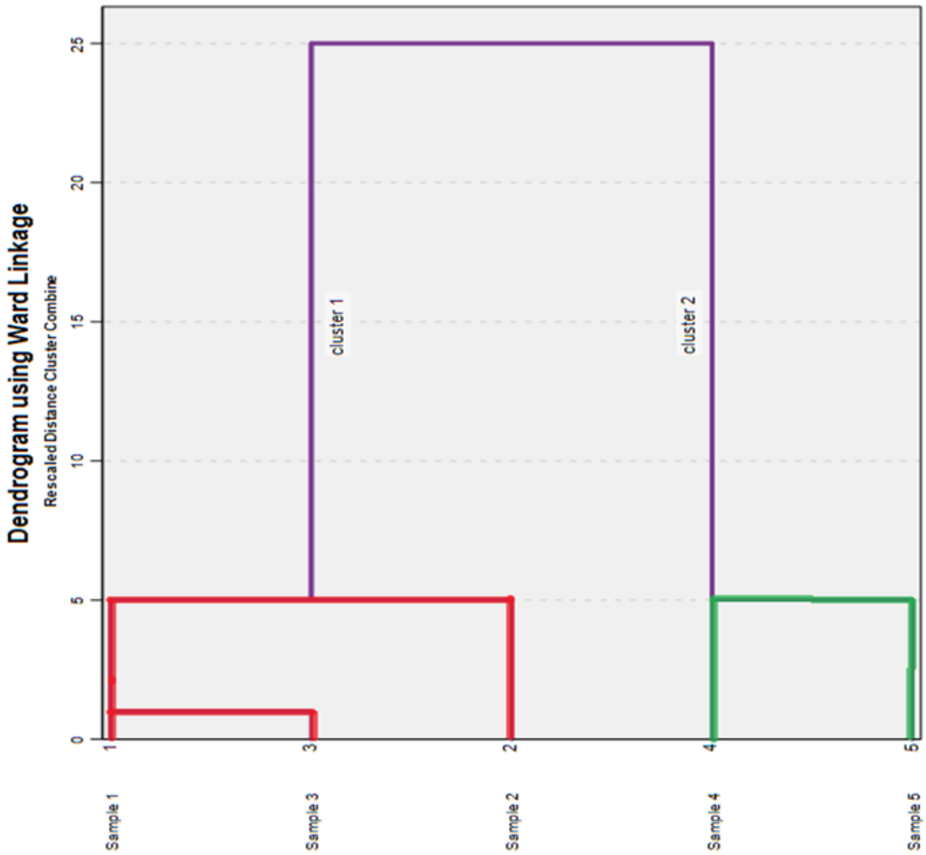


Fig 4.2: Dendrogram for Surface water in October month

Dendrograms provide a visual representation of the clustering process, showing how similar sites are grouped into clusters based on distance metrics. In this case, Cluster 1 includes Sample 1, 2 and 3 whereas Cluster 2 includes Sample 4 and 5.

4.1.2 Results showing Surface water analysis using HCA for November month

Table 4.4: Table showing Processing Summary

Case Processing Summary					
Cases					
Valid		Missing		Total	
N	Percent	N	Percent	N	Percent
5	100.0%	0	0.0%	5	100.0%
a. Squared Euclidean Distance used					
b. Ward Linkage					

Table 4.5: Similarity Matrix for Surface water in November month

Proximity Matrix					
Case	Squared Euclidean Distance				
	1: Tirap River Up-Stream near Lekhapani Bridge	2: Tirap River Down-Stream (near Molong bangali gaon 1)	3: Buri Dehing River Up-Stream (Near Manmae Maichang Gaon)	4: Buri Dehing Down-Stream (Near Circuit House)	5: Mixing Zone of Buri Dehing and Ledo Pani Mallah (Near Ledo Namdang Gaon)
1: Sample 1	.000	6.736	16.848	13.410	60.397
2: Sample 2	6.736	.000	9.911	13.110	58.170
3: Sample 3	16.848	9.911	.000	20.761	88.316
4: Sample 4	13.410	13.110	20.761	.000	52.342
5: Sample 5	60.397	58.170	88.316	52.342	.000

The degree of similarity (or dissimilarity) between different sampling sites using squared Euclidean distance is quantified in this table. The most dissimilar samples are Sample 3 and 5.

Table 4.6: Ward linkage for Surface water in November month.

Agglomeration Schedule						
Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	1	2	3.368	0	0	2
2	1	4	11.085	1	0	3
3	1	3	20.194	2	0	4
4	1	5	68.000	3	0	0

The table highlights the step-by-step process of forming clusters, showing how individual or grouped samples combine based on Ward's method. This table consist of four stages with highest coefficient for stage 4.

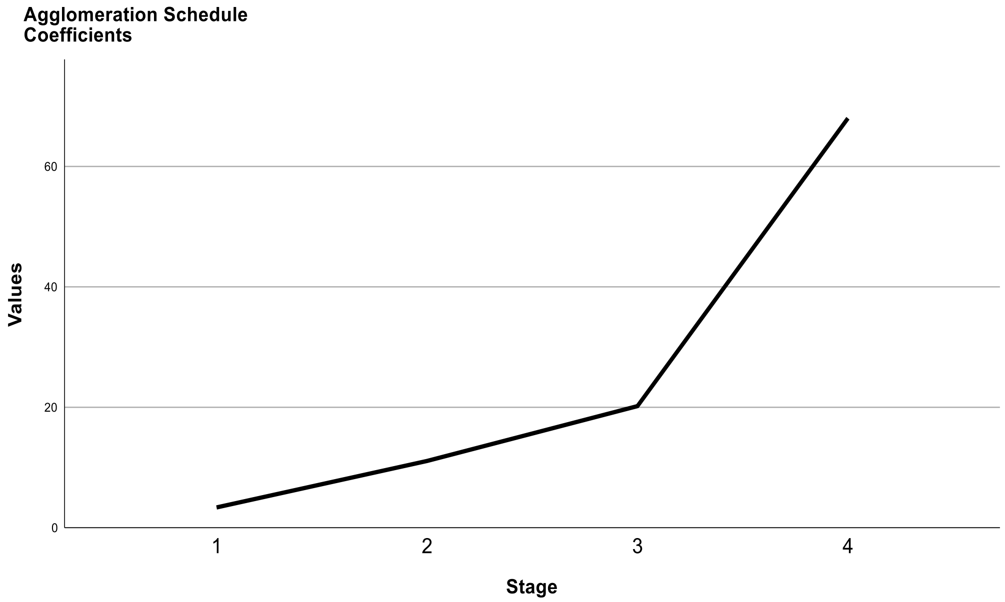


Fig 4.3: Icicle plots for Surface water in November month

The plots of agglomeration schedule coefficients (icicle plots) indicate the strength of cluster formation at each stage, with sharp increase in coefficients suggesting optimal cluster points.

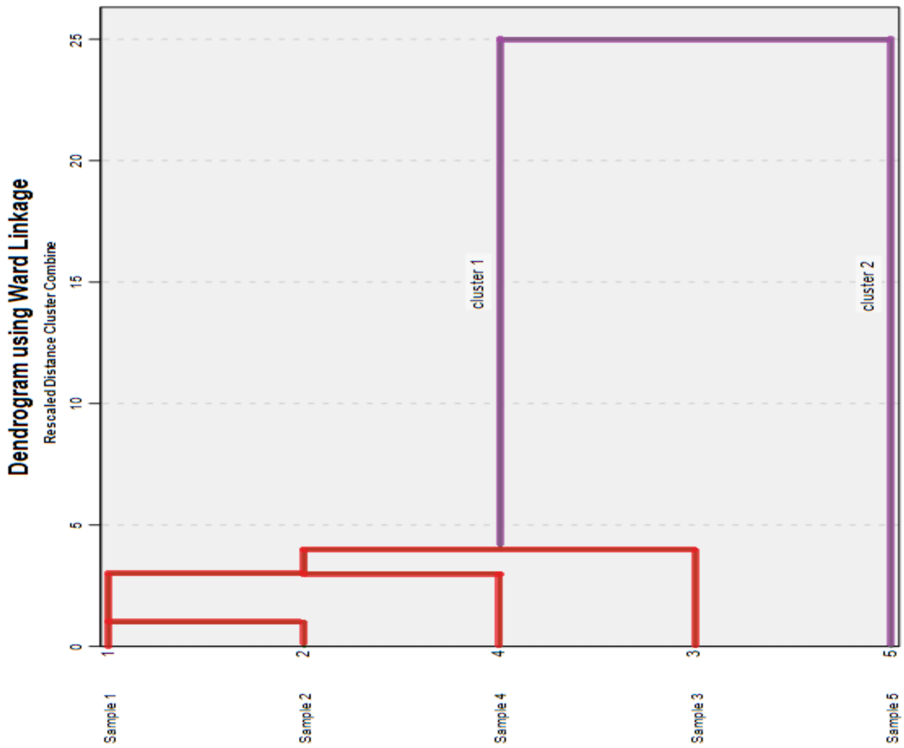


Fig 4.4: Dendrogram for Surface water in November month

Dendrograms provide a visual representation of the clustering process, showing how similar sites are grouped into clusters based on distance metrics. In this case, Cluster 1 includes Sample 1, 2, 3 and 4 whereas Cluster 2 includes only Sample 5.

4.1.3 Results showing Surface water analysis using HCA for December month

Table 4.7: Table showing Processing Summary

Case Processing Summary					
Cases					
Valid		Missing		Total	
N	Percent	N	Percent	N	Percent
5	100.0%	0	0.0%	5	100.0%
a. Squared Euclidean Distance used					
b. Ward Linkage					

Table 4.8: Similarity Matrix for Surface water in December month

Proximity Matrix					
Case	Squared Euclidean Distance				
	1: Tirap River Up-Stream near Lekhapani Bridge	2: Tirap River Down-Stream (near Molong bangali gaon 1)	3: Buri Dehing River Up-Stream (Near Manmae Maichang Gaon)	4: Buri Dehing Down-Stream (Near Circuit House)	5: Mixing Zone of Buri Dehing and Ledo Pani Mallah (Near Ledo Namdang Gaon)
1: Sample 1	.000	6.580	6.352	27.842	53.968
2: Sample 2	6.580	.000	8.444	29.640	56.694
3: Sample 3	6.352	8.444	.000	17.612	59.216
4: Sample 4	27.842	29.640	17.612	.000	73.651
5: Sample 5	53.968	56.694	59.216	73.651	.000

The degree of similarity (or dissimilarity) between different sampling sites using squared Euclidean distance is quantified in this table. The most dissimilar samples are Sample 4 and 5.

Table 4.9: Ward linkage for Surface water in December month

Agglomeration Schedule						
Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	1	3	3.176	0	0	2
2	1	2	7.125	1	0	3
3	1	4	24.118	2	0	4
4	1	5	68.000	3	0	0

The table highlights the step-by-step process of forming clusters, showing how individual or grouped samples combine based on Ward's method. This table consist of four stages with highest coefficient for stage 4.

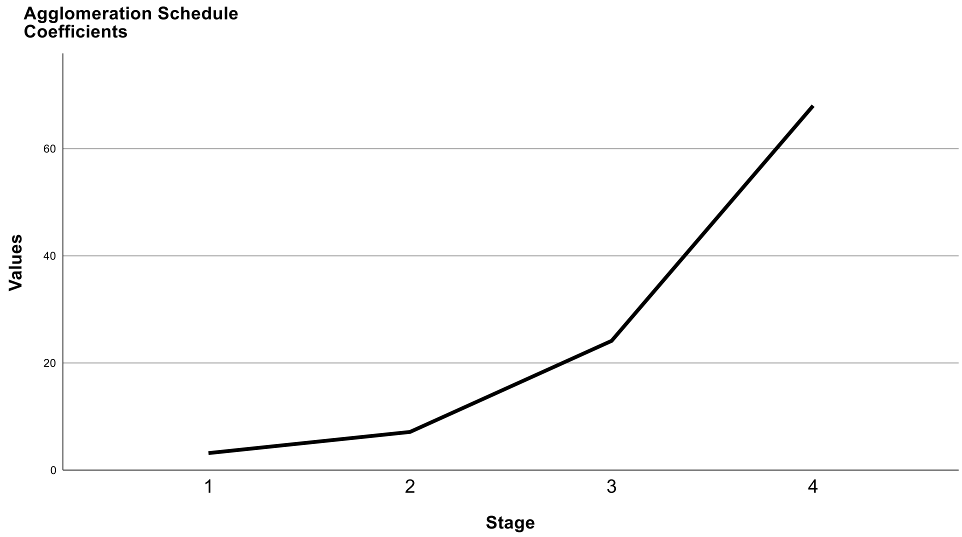


Fig 4.5: Icicle plots for Surface water in December month

The plots of agglomeration schedule coefficients (icicle plots) indicate the strength of cluster formation at each stage, with sharp increase in coefficients suggesting optimal cluster points.

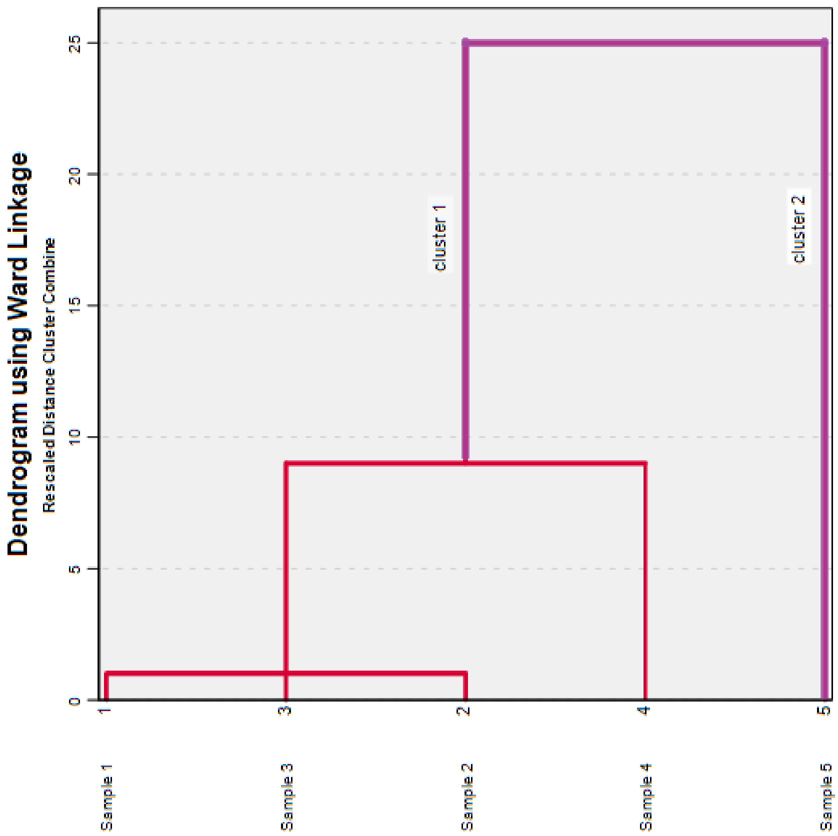


Fig 4.6: Dendrogram for Surface water in December month

Dendrograms provide a visual representation of the clustering process, showing how similar sites are grouped into clusters based on distance metrics. In this case, Cluster 1 includes Sample 1, 2, 3 and 4 whereas Cluster 2 includes Sample 5.

4.2 Hierarchical Cluster Analysis for Effluent water:

4.2.1 Results showing Effluent water analysis using HCA for October month

Table 4.10: Table showing Processing Summary

Case Processing Summary					
Cases					
Valid		Missing		Total	
N	Percent	N	Percent	N	Percent
3	100.0%	0	0.0%	3	100.0%
a. Squared Euclidean Distance used					
b. Ward's linkage					

Table 4.11: Similarity matrix for Effluent water in October month

Proximity Matrix			
Case	Squared Euclidean Distance		
	1: Tirap East Near Sedimentation Tank	2: Tirap West Near Discharge Point	3: Water-Sedimentation Tank of Tirap OCP
1: Sample 1	.000	42.502	35.612
2: Sample 2	42.502	.000	41.886
3: Sample 3	35.612	41.886	.000

The degree of similarity (or dissimilarity) between different sampling sites using squared Euclidean distance is quantified in this table. The most dissimilar samples are Sample 1 and 2.

Table 4.12: Ward linkage for Effluent water in October month

Agglomeration Schedule						
Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	1	3	17.806	0	0	2
2	1	2	40.000	1	0	0

The table highlights the step-by-step process of forming clusters, showing how individual or grouped samples combine based on Ward's method. This table consist of four stages with highest coefficient for stage 4.

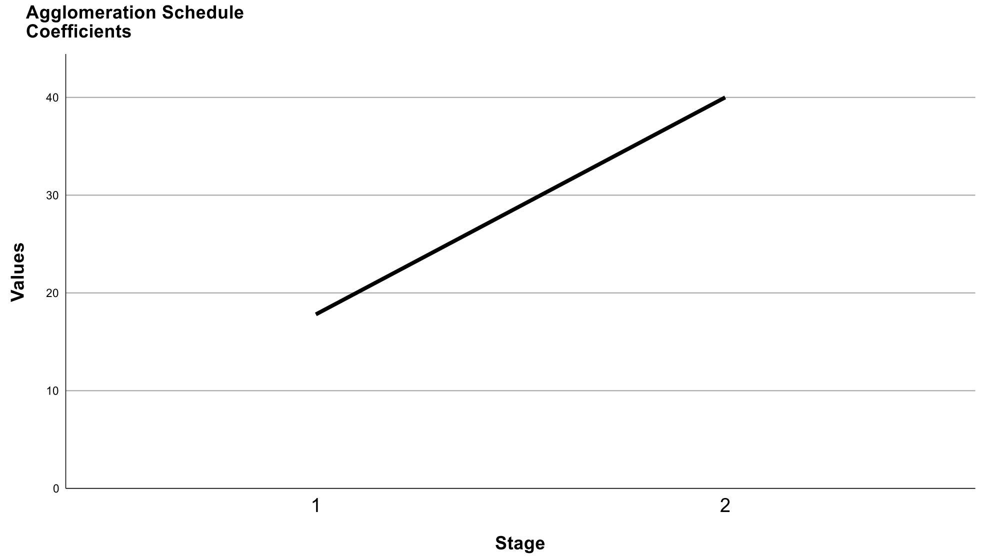


Fig 4.7: Icicle plots for Effluent water in October month

The plots of agglomeration schedule coefficients (icicle plots) indicate the strength of cluster formation at each stage, with sharp increase in coefficients suggesting optimal cluster points.

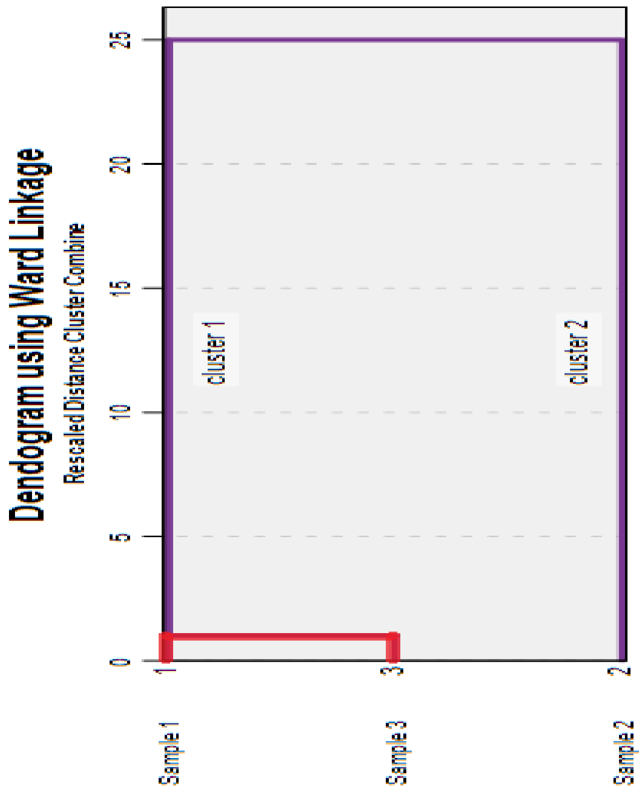


Fig 4.8: Dendrogram for Effluent water in October month

Dendrograms provide a visual representation of the clustering process, showing how similar sites are grouped into clusters based on distance metrics. In this case, Cluster 1 includes Sample 1 and 3 whereas Cluster 2 includes Sample 2.

4.2.2 Results showing Effluent water analysis using HCA for November month

Table 4.13: Table showing Processing Summary

Case Processing Summary					
Cases					
Valid		Missing		Total	
N	Percent	N	Percent	N	Percent
3	100.0%	0	0.0%	3	100.0%
a. Squared Euclidean Distance used					
b. Ward Linkage					

Table 4.14: Similarity matrix for Effluent water in November month

Proximity Matrix			
Case	Squared Euclidean Distance		
	1: Tirap East Near Sedimentation Tank	2: Tirap West Near Discharge Point	3: Water-Sedimentation Tank of Tirap OCP
1: Sample 1	.000	22.612	39.024
2: Sample 2	22.612	.000	40.365
3: Sample 3	39.024	40.365	.000

The degree of similarity (or dissimilarity) between different sampling sites using squared Euclidean distance is quantified in this table. The most dissimilar samples are Sample 2 and 3.

Table 4.15: Ward linkage for Effluent water in November month

Agglomeration Schedule						
Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	1	2	11.306	0	0	2
2	1	3	34.000	1	0	0

The table highlights the step-by-step process of forming clusters, showing how individual or grouped samples combine based on Ward's method. This table consist of four stages with highest coefficient for stage 4.

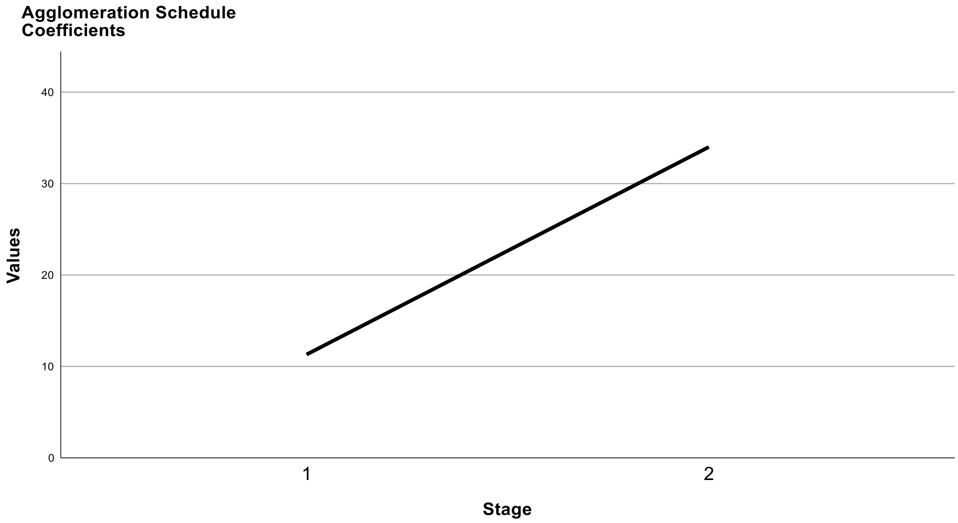


Fig 4.9: Icicle plots for Effluent water in November month

The plots of agglomeration schedule coefficients (icicle plots) indicate the strength of cluster formation at each stage, with sharp increase in coefficients suggesting optimal cluster points.

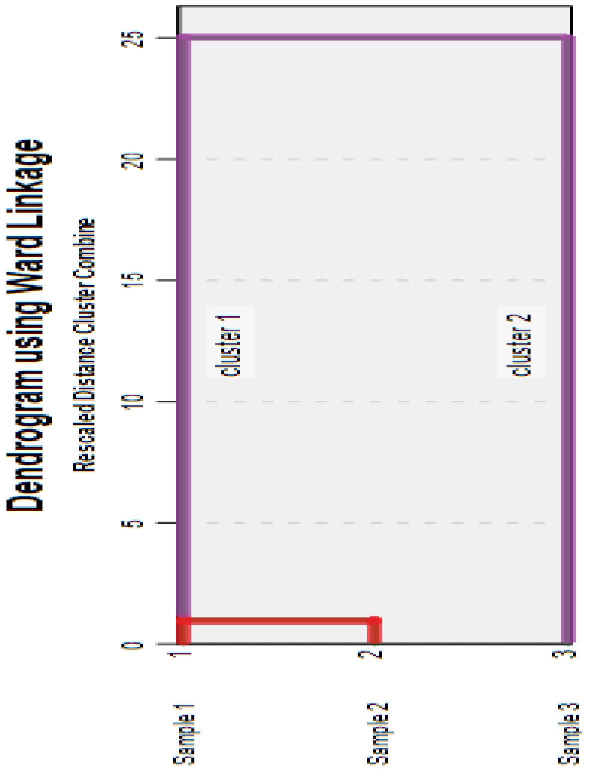


Fig 4.10: Dendrogram for Effluent water in November month

Dendrograms provide a visual representation of the clustering process, showing how similar sites are grouped into clusters based on distance metrics. In this case, Cluster 1 includes Sample 1 and 2 whereas Cluster 2 includes Sample 3.

4.2.3 Results showing Effluent water analysis using HCA for December month

Table 4.16: Table showing Processing Summary

Case Processing Summary					
Cases					
Valid		Missing		Total	
N	Percent	N	Percent	N	Percent
3	100.0%	0	0.0%	3	100.0%
a. Squared Euclidean Distance used					
b. Ward Linkage					

Table 4.17: Similarity matrix for Effluent water in December month

Proximity Matrix			
Case	Squared Euclidean Distance		
	1: Tirap East Near Sedimentation Tank	2: Tirap West Near Discharge Point	3: Water-Sedimentation Tank of Tirap OCP
1: Sample 1	.000	29.406	24.222
2: Sample 2	29.406	.000	36.372
3: Sample 3	24.222	36.372	.000

The degree of similarity (or dissimilarity) between different sampling sites using squared Euclidean distance is quantified in this table. The most dissimilar samples are Sample 2 and 3.

Table 4.18: Ward linkage for Effluent water in December month

Agglomeration Schedule						
Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	1	3	12.111	0	0	2
2	1	2	30.000	1	0	0

The table highlights the step-by-step process of forming clusters, showing how individual or grouped samples combine based on Ward's method. This table consist of four stages with highest coefficient for stage 4.

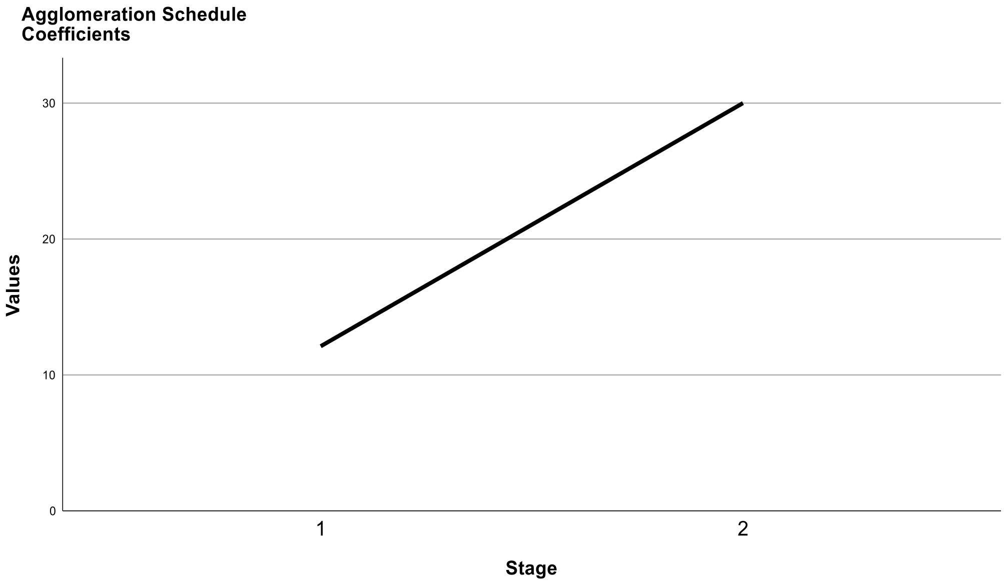


Fig 4.11: Icicle plots for Effluent water in December month

The plots of agglomeration schedule coefficients (icicle plots) indicate the strength of cluster formation at each stage, with sharp increase in coefficients suggesting optimal cluster points.

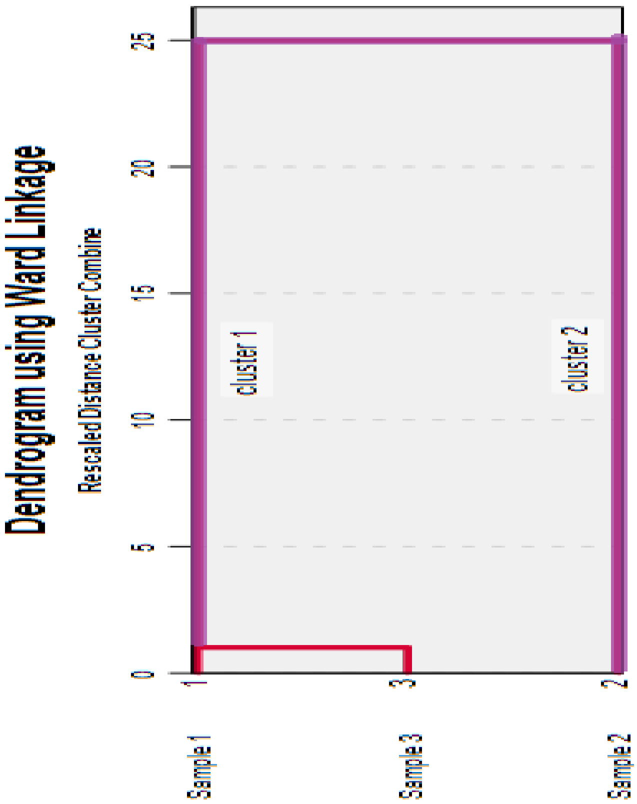


Fig 4.12: Dendrogram for Effluent water in December month

Dendrograms provide a visual representation of the clustering process, showing how similar sites are grouped into clusters based on distance metrics. In this case, Cluster 1 includes Sample 1 and 3 whereas Cluster 2 includes Sample 2.

4.3 Hierarchical Cluster Analysis for Drinking water:

4.3.1 Results showing Drinking water analysis using HCA for October month

Table 4.19: Table showing Processing Summary

Case Processing Summary					
Cases					
Valid		Missing		Total	
N	Percent	N	Percent	N	Percent
6	100.0%	0	0.0%	6	100.0%
a. Squared Euclidean Distance used					
b. Ward Linkage					

Table 4.20: Similarity matrix for Drinking water in October month

Proximity Matrix						
Case	Squared Euclidean Distance					
	1: Jagun Market	2: Tipong near Punjab National Bank	3: Saliki Gaon	4: Tirap Core Zone	5: Near Ledo College	6: Segunbari, Margherita
1: Sample 1	.000	9.938	37.811	19.091	11.202	6.931
2: Sample 2	9.938	.000	47.365	4.446	8.776	13.519
3: Sample 3	37.811	47.365	.000	55.284	47.725	44.436
4: Sample 4	19.091	4.446	55.284	.000	14.335	25.363
5: Sample 5	11.202	8.776	47.725	14.335	.000	13.778
6: Sample 6	6.931	13.519	44.436	25.363	13.778	.000

The degree of similarity (or dissimilarity) between different sampling sites using squared Euclidean distance is quantified in this table. The most dissimilar samples are Sample 3 and 4.

Table 4.21: Ward linkage for Drinking water in October month

Agglomeration Schedule						
Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	2	4	2.223	0	0	3
2	1	6	5.689	0	0	4
3	2	5	12.651	1	0	4
4	1	2	25.476	2	3	5
5	1	3	60.000	4	0	0

The table highlights the step-by-step process of forming clusters, showing how individual or grouped samples combine based on Ward's method. This table consist of four stages with highest coefficient for stage 4.

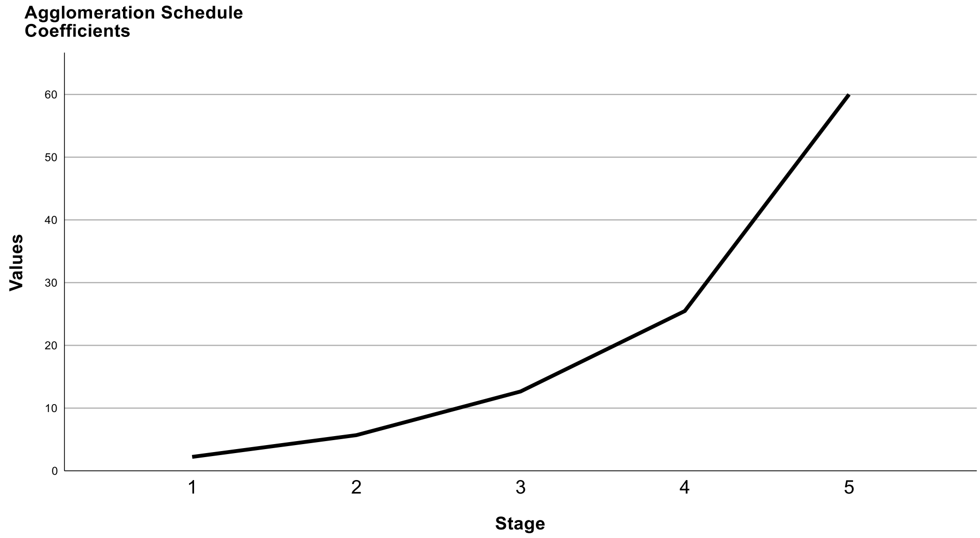


Fig 4.13: Icicle plots for Drinking water in October month

The plots of agglomeration schedule coefficients (icicle plots) indicate the strength of cluster formation at each stage, with sharp increase in coefficients suggesting optimal cluster points.

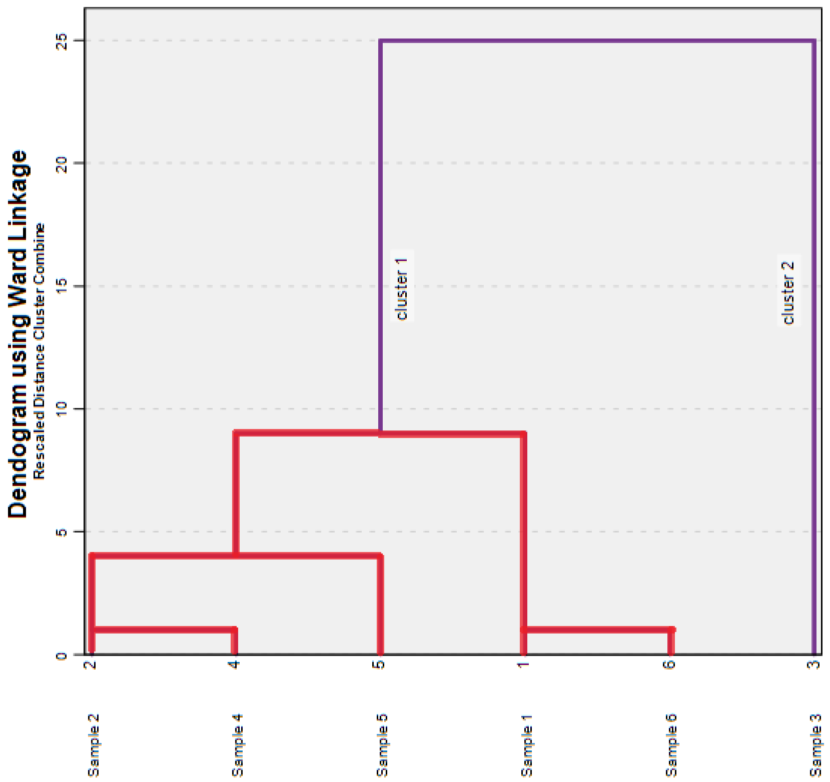


Fig 4.14: Dendrogram for Drinking water in October month

Dendrograms provide a visual representation of the clustering process, showing how similar sites are grouped into clusters based on distance metrics. In this case, Cluster 1 includes Sample 1, 2, 4, 5 and 6 whereas Cluster 2 includes Sample 3.

4.3.2 Results showing Drinking water analysis using HCA for November month

Table 4.22: Table showing Processing Summary

Case Processing Summary					
Cases					
Valid		Missing		Total	
N	Percent	N	Percent	N	Percent
6	100.0%	0	0.0%	6	100.0%
a. Squared Euclidean Distance used					
b. Ward Linkage					

Table 4.23: Similarity matrix for Drinking water in November month

Proximity Matrix						
Case	Squared Euclidean Distance					
	1: Jagun Market	2: Tipong near Punjab National Bank	3: Saliki Gaon	4: Tirap Core Zone	5: Near Ledo College	6: Segunbari, Margherita
1: Sample 1	.000	36.179	20.173	36.250	33.970	32.028
2: Sample 2	36.179	.000	14.927	24.976	20.944	53.267
3: Sample 3	20.173	14.927	.000	15.459	22.421	43.159
4: Sample 4	36.250	24.976	15.459	.000	27.251	54.690
5: Sample 5	33.970	20.944	22.421	27.251	.000	44.307
6: Sample 6	32.028	53.267	43.159	54.690	44.307	.000

The degree of similarity (or dissimilarity) between different sampling sites using squared Euclidean distance is quantified in this table. The most dissimilar samples are Sample 4 and 6.

Table 4.24: Ward linkage for Drinking water in November month

Agglomeration Schedule						
Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	2	3	7.463	0	0	2
2	2	4	18.454	1	0	3
3	2	5	31.494	2	0	5
4	1	6	47.509	0	0	5
5	1	2	80.000	4	3	0

The table highlights the step-by-step process of forming clusters, showing how individual or grouped samples combine based on Ward's method. This table consist of four stages with highest coefficient for stage 4.

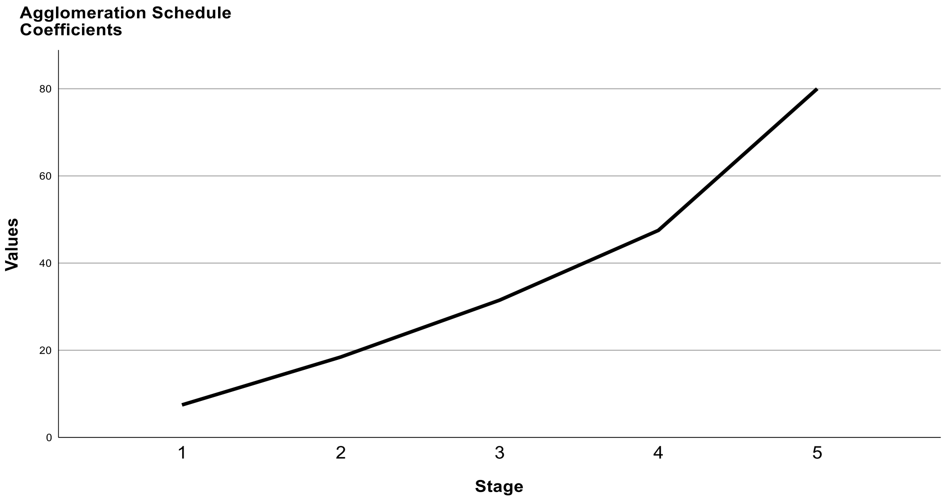


Fig 4.15: Icicle plots for Drinking water in November month

The plots of agglomeration schedule coefficients (icicle plots) indicate the strength of cluster formation at each stage, with sharp increase in coefficients suggesting optimal cluster points.

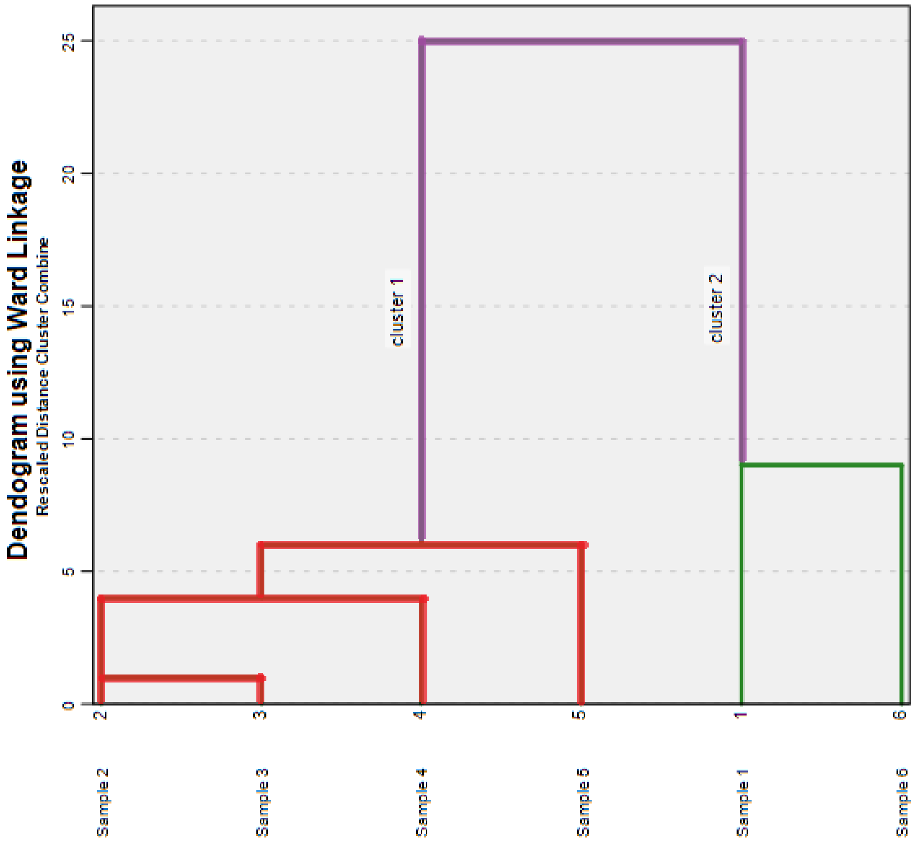


Fig 4.16: Dendrogram for Drinking water in November month

Dendrograms provide a visual representation of the clustering process, showing how similar sites are grouped into clusters based on distance metrics. In this case, Cluster 1 includes Sample 2, 3, 4 and 5 whereas Cluster 2 includes Sample 1 and 6.

4.3.3 Results showing Drinking water analysis using HCA for December month

Table 4.25: Table showing Processing Summary

Case Processing Summary					
Cases					
Valid		Missing		Total	
N	Percent	N	Percent	N	Percent
6	100.0%	0	0.0%	6	100.0%
a. Squared Euclidean Distance used					
b. Ward Linkage					

Table 4.26: Similarity matrix for Drinking water in December month

Proximity Matrix						
Case	Squared Euclidean Distance					
	1: Jagun Market	2: Tipong near Punjab National Bank	3: Saliki Gaon	4: Tirap Core Zone	5: Near Ledo College	6: Segunbari, Margherita
1: Sample 1	.000	40.765	17.955	34.063	25.796	26.376
2: Sample 2	40.765	.000	25.017	27.953	40.503	64.856
3: Sample 3	17.955	25.017	.000	16.869	18.898	24.928
4: Sample 4	34.063	27.953	16.869	.000	18.771	38.118
5: Sample 5	25.796	40.503	18.898	18.771	.000	29.131
6: Sample 6	26.376	64.856	24.928	38.118	29.131	.000

The degree of similarity (or dissimilarity) between different sampling sites using squared Euclidean distance is quantified in this table. The most dissimilar samples are Sample 2 and 6.

Table 4.27: Ward linkage for Drinking water in December month

Agglomeration Schedule						
Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	3	4	8.435	0	0	2
2	3	5	18.179	1	0	4
3	1	6	31.367	0	0	4
4	1	3	50.181	3	2	5
5	1	2	75.000	4	0	0

The table highlights the step-by-step process of forming clusters, showing how individual or grouped samples combine based on Ward's method. This table consist of four stages with highest coefficient for stage 4.

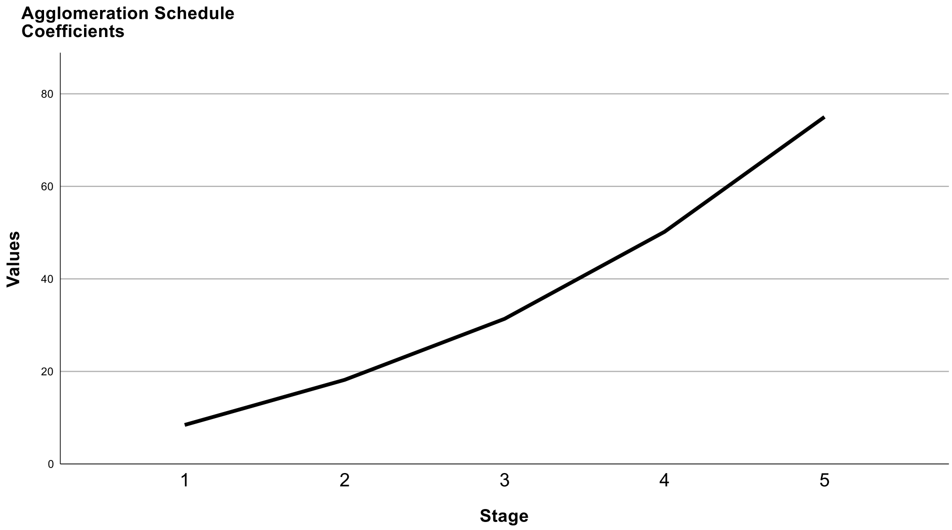


Fig 4.17: Icicle plots for Drinking water in December month

The plots of agglomeration schedule coefficients (icicle plots) indicate the strength of cluster formation at each stage, with sharp increase in coefficients suggesting optimal cluster points.

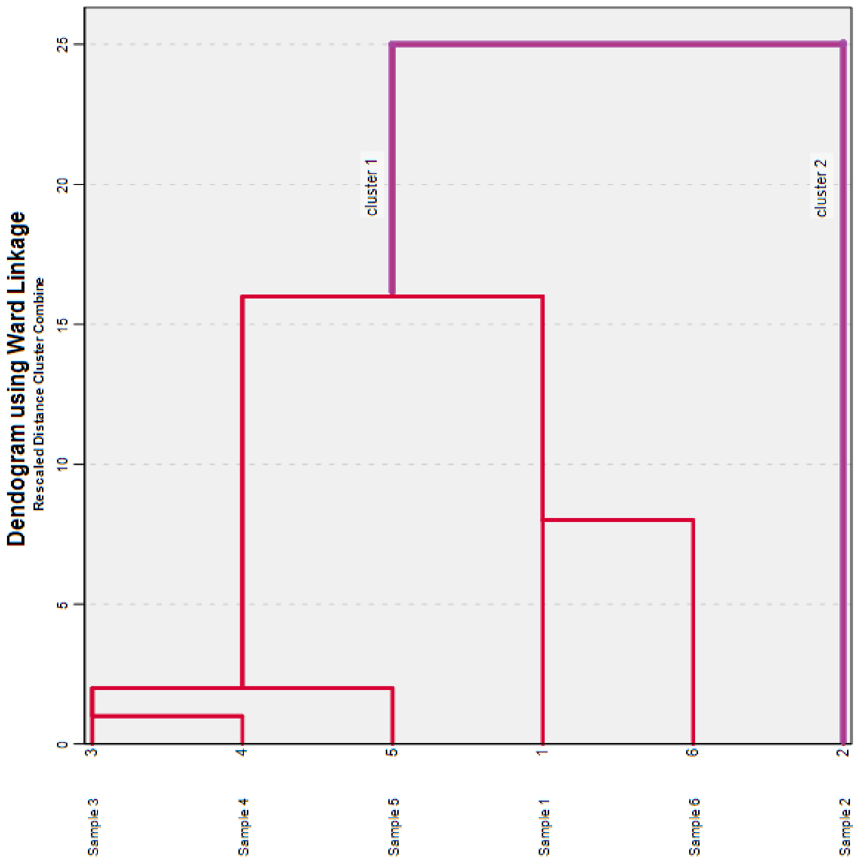


Fig 4.18: Dendrogram for Drinking water in December month

Dendrograms provide a visual representation of the clustering process, showing how similar sites are grouped into clusters based on distance metrics. In this case, Cluster 1 includes Sample 1, 3, 4, 5 and 6 whereas Cluster 2 includes Sample 2.

Discussions

From the dendrograms of the Hierarchical Cluster Analysis (HCA), the variation between the clusters provides valuable insights into the pollution levels. A cluster exhibiting higher variation indicates higher pollution levels compared to the other cluster.

4.3.4 Surface Water:

- **October:** The dendrogram shows that Cluster 1 includes Samples 1, 2, and 3, while Cluster 2 includes Samples 4 and 5. The samples exhibit similar properties in the tested parameters within each of the two clusters. However, both clusters are showing similar overall variations as depicted in the dendrogram. So, we can say that both the clusters have equal pollution.
- **November:** Cluster 1 comprises Samples 1, 2, 3, and 4, while Cluster 2 includes only Sample 5. Samples in Cluster 1(sample 1, 2, 4 and 4) exhibit similar properties in the tested parameters. The dendrogram shows that Cluster 2 have higher variation than Cluster 1, so we can say that Cluster 2 (Sample 5) have higher pollution than Cluster 1.
- **December:** Cluster 1 consists of Samples 1, 2, 3, and 4, while Cluster 2 includes only Sample 5. Samples 1, 2, 3 and 4 exhibit similar properties within Cluster 1. The dendrogram shows that Cluster 2 have higher variation than Cluster 1, so we can say that Cluster 2 (Sample 5) again have higher pollution than Cluster 1.

4.3.5 Effluent Water:

- **October:** The dendrogram shows that Cluster 1 includes Samples 1 and 3, while Cluster 2 includes only Samples 2. Samples 1 and 3 exhibit similar properties within Cluster 1. Cluster 2 (Sample 2) exhibits higher variation compared to Cluster 1. So, we can say that Cluster 2 (Sample 2) have higher pollution than Cluster 1.
- **November:** Cluster 1 consists of Samples 1 and 2, while Cluster 2 includes only Sample 3. Samples 1 and 2 exhibit similar properties within Cluster 1. The dendrogram shows that Cluster 2 have higher variation than Cluster 1, so we can say that Cluster 2 (Sample 3) have higher pollution than Cluster 1.
- **December:** Cluster 1 consists of Samples 1 and 3, while Cluster 2 includes only Sample 2. Samples 1 and 3 exhibit similar properties within Cluster 1. The dendrogram shows that

Cluster 2 have higher variation than Cluster 1, so we can say that Cluster 2 (Sample 2) have higher pollution than Cluster 1.

4.3.6 Drinking Water:

- **October:** The dendrogram shows that Cluster 1 include Samples 1, 2, 4, 5 and 6, while Cluster 2 include only Sample 3. Samples 1, 2, 4, 5 and 6 exhibit similar properties within Cluster 1. Cluster 2 (Sample 3) demonstrates higher variation compared to Cluster 1. So, we can say that Cluster 2 (Sample 3) have higher pollution than Cluster 1.
- **November:** The dendrogram shows that Cluster 1 include Samples 2, 3, 4 and 5, while Cluster 2 include Samples 1 and 6. Samples 2, 3, 4 and 5 exhibit similar properties within Cluster 1 and Samples 1 and 6 exhibit similar properties within Cluster 2. Cluster 2 (Sample 1 and 6) demonstrates higher variation compared to Cluster 1. So, we can say that Cluster 2 (Sample 1 and 6) have higher pollution than Cluster 1.
- **December:** The dendrogram shows that Cluster 1 include Samples 1, 3, 4, 5 and 6, while Cluster 2 include only Sample 2. Samples 1, 3, 4, 5 and 6 exhibit similar properties within Cluster 1. Cluster 2 (Sample 2) demonstrates higher variation compared to Cluster 1. So, we can say that Cluster 2 (Sample 2) again have higher pollution than Cluster 1.

The analysis highlights monthly variations in pollution levels across the samples collected from different locations. The clustering patterns reveal the relative levels of pollution among the samples, with the clusters having higher variation indicating samples with greater pollution levels based on the tested parameters. The permissible limits of various parameters for Surface water, Effluent water and Drinking water were referred from IS 2296, Standards for Coal Mines and IS 10500 – 2012 respectively.

From the above discussion, it can be analysed that among Surface water samples, Sample 5 remained polluted in all the three months. Fig. 3.2 represents that Sample 5 belongs to the location Mixing Zone of Buri Dehing and Ledo Pani Mallah (Near Ledo Namdang Gaon) which are at the downstream to the Tirap core Zone. Among Effluent water Samples, Sample 2 is found to be polluted for October and December month whereas Sample 3 is found to be polluted in the November month. Fig. 3.3 represents Sample 2 belongs to the location Tirap West Near Discharge Point and Sample 3 belongs to Treated Effluent Water from the Sedimentation Tank of Tirap OCP. Both the locations are at the downstream to the Tirap core Zone. Among Drinking water Samples, Sample 3 is found to be polluted for October month, Sample 1 and 6 found to be polluted for

November month whereas Sample 2 found to be polluted in the December month. Fig. 3.4 represents Sample 1 belongs to the location Jagun Market, Sample 2 belongs to Tipong near PNB, Sample 3 belongs to Saliki gaon and Sample 6 belongs to Segunbari, Margherita. Locations of Sample 1, 2 and 3 are at the upstream side whereas Sample 6 at the downstream side of Tirap Core Zone. Based on this analysis, it can be inferred that the pollution in surface and effluent water is likely attributed to mine leaching, while the contamination observed in drinking water indicates potential groundwater pollution.

Kataki and Borah (2024) has utilized similar methods (HCA) to address pollution sources and to assess the water's suitability for aquatic life of Physio-chemical parameters in the Deepor Beel, located in Kamrup, Assam. The study takes sampling from nine fixed locations across four seasons (autumn, winter, spring, summer), providing a seasonal perspective over a full year. Hierarchical Cluster Analysis groups, the sampling sites into two clusters for each season based on similar sites. The spatial pattern shows that some of the sites had the lowest level of pollution while other sampling sites had higher levels of pollution. It investigates pollution influenced by mixed-use land activities (agriculture, urbanization, industrial discharge). It analyses a broader range of parameters, including salinity, biochemical oxygen demand (BOD), turbidity, electrical conductivity (EC), and heavy metals (e.g., lead, iron). High nitrate and lead concentrations were noted due to domestic and industrial effluents in Autumn and Summer seasons. Whereas higher dissolved oxygen (DO) levels and reduced BOD indicated improved water quality during cooler months i.e. Winter and Spring. Specific sites near residential areas and industrial zones consistently exhibited higher contamination (e.g., salinity, EC). Lead concentrations exceeded permissible limits, particularly during autumn (from urban and industrial runoff). Salinity, turbidity, and EC peaked during summer, reflecting concentration effects due to reduced dilution. Seasonal variations in water quality reflect the interaction of natural and anthropogenic factors. The study suggests urgent management interventions to control agricultural runoff and industrial effluents.

5 CONCLUSION

5.1 OVERVIEW

This study successfully applied Hierarchical Cluster Analysis (HCA) to assess the quality of surface, effluent, and drinking water in and around the Tirap Open Cast Project (OCP), located in the Makum Coalfield of Tinsukia District, Assam. The analysis revealed significant seasonal and spatial variations in water quality parameters, highlighting the impacts of mining activities on nearby water resources.

5.2 OUTCOMES

Outcomes demonstrate that water samples from certain clusters exhibited comparatively higher levels of pollution, indicating the need for targeted interventions. The results underscore the critical role of **Multivariate Statistical Analysis (MSA)** in identifying pollution hotspots and guiding decision-making processes for sustainable water management. From these analyses, the following conclusions are drawn:

1. **Surface Water**: The surface water samples found to be comparatively polluted are located at the downstream of the Tirap Core Zone. The parameters in higher concentrations at these locations include **Total Hardness**, **Calcium**, **Chloride**, **Magnesium**, **Total Dissolved Solids (TDS)**, **Sulphate**, and **Fluoride**.
2. **Effluent Water**: The effluent water samples having higher pollution are from the locations Tirap West Near Discharge Point and Treated Effluent Water from the Sedimentation Tank of Tirap OCP across all three months which are at the downstream of the Tirap Core Zone. The parameters in higher concentrations at these locations include **Total Kjeldahl Nitrogen (TKN)**, **Chemical Oxygen Demand (COD)**, **Biochemical Oxygen Demand (BOD)**, **Chloride**, and **Turbidity**.
3. **Drinking Water**: The drinking water samples found to be comparatively polluted are from the locations Jagun Market, Tipong Near Punjab National Bank, Saliki Gaon, and Segunbari, Margherita across all three months. The parameters in higher concentrations at these locations include **Total Hardness**, **Total Alkalinity**, **TDS**, **Chloride**, **Sulphate**, and **Sodium**.

It can be concluded that the pollution in surface and effluent water is likely caused by mine leaching, while the contamination observed in drinking water suggests potential groundwater pollution. However, it is important to note that all collected parameter values fall within the permissible limits outlined in the respective standard codes, indicating that the water is only relatively polluted. While urgent interventions may not be necessary, implementing effective mitigation strategies would be beneficial.

5.3 SCOPE FOR FUTURE RESEARCH

This study provides valuable insights; however, the limitations related to temporal and spatial scope, parameter selection, and external influences highlight the need for more comprehensive future research. Combining effective mitigation strategies with robust monitoring programs is essential to minimize the environmental impacts of mining activities and safeguard regional water resources.

6 REFERENCES

1. Ammar Tiri, Noureddine Lahbari (2015) "Assessment of the quality of water by hierarchical cluster and variance analyses of the Koudiat Medouar Watershed, East Algeria", Appl Water Sci (2017) 7:4197–4206.
2. Ammar Salman Dawood (2017) "Using Multivariate Statistical Methods for The Assessment of The Surface Water Quality for A River: A Case Study", International Journal of Civil Engineering and Technology (IJCIET), Volume 8, Issue 10, October 2017, pp. 588–597.
3. Barrie, S.K. Agodzo, R. Frazer-Williams, E. Awuah, E. Bessah (2023) "A multivariate statistical approach and water quality index for water quality assessment for the Rokel river in Sierra Leone", Heliyon 9 (2023) e16196, 2405-8440.
4. Hydrogeology Department Exploration Division CMPDI, Ranchi "Comprehensive Report on Groundwater condition for Tirap OCP, NEC", March, 2022.
5. Kuldeep Singh Rautela, Dilip Kumar, Bhishm Singh Khatia, Amit Rawat and Amit Kumar Dubey (2023) "Multivariate statistical analysis to assess the surface water quality of a snow and glacierfed river: A case from Alaknanda River basin", Water Science & Technology Vol 88 No 11, 2873, doi: 10.2166/wst.2023.349.
6. Lakshana Kataki, Dr. Triptimoni Borah (2024) "Statistical Assessment of Water Quality Parameters for Pollution Source Identification in Deepor Beel Area, Guwahati, Assam", International Journal for Multidisciplinary Research (IJFMR), E-ISSN: 2582-2160.
7. Standards For Coal Mines, MOEF, GSR-742(E), dt. 25.09.2000 and GSR-801 (E), EPA, 1986, dt. 31.12.1993.