Application of Discrete Mathematics in Groundwater Quality Assessment: A Case Study from Rahata Tehsil, Ahilyanagar District, Maharashtra, India

Aruna Santosh Zine[1*], Kotkar Bharti Amol[2]

[1*]Post Graduate, Department of Environment Science And Research Centre, Padmashri Vikhe Patil College of Arts, Science & Commerce, Pravaranagar, Maharashtra, India, Email ID: aruna.thorat@gmail.com,
ORCID: https://orcid.org/0009-0005-1390-4353
[2]Department of Physics, Padmashri Vikhe Patil College of Arts Science & Commerce, Pravaranagar, Maharashtra, India, 413713. Email ID: bhartipvp21@gmail.com

Abstract
Discrete mathematics offers powerful analytical tools for modeling and evaluating environmental data, particularly in scenarios involving classification, relational analysis, and pattern enumeration. This study demonstrates the application of set theory, graph theory, and combinatorics to assess groundwater quality parameters from 34 samples (17 tube wells and 17 open wells) collected across 17 villages in Rahata Tehsil, Ahmednagar District, and Maharashtra. Key hydro chemical parameters analyzed include pH, electrical conductivity (EC), major ions (Ca^{2+}, Mg^{2+}, Na^{+}, K^{+}, HCO_3^{-}, Cl^{-}, SO_4^{2-}), total dissolved solids (TDS), and total hardness. Data discretization enables set-based classification of quality levels, graph-theoretic analysis reveals parameter interdependencies through correlation networks, and combinatorial methods quantify multi-parameter exceedance patterns against Bureau of Indian Standards (BIS IS 10500:2012) drinking water specifications. Results demonstrate strong graph connectivity among salinity-related parameters (EC, TDS, Mg^{2+}, Na^{+}) with Pearson correlation coefficients exceeding 0.8, indicating geogenic linkages characteristic of basaltic aquifer weathering. Thirty samples (88%) exceeded TDS acceptability limits of 500 mg/L, while combinatorial analysis identified two samples with extreme multi-parameter violations. These discrete mathematical approaches provide a structured, computationally efficient framework for identifying groundwater quality hotspots and supporting sustainable water resource management in semi-arid basaltic terrains.

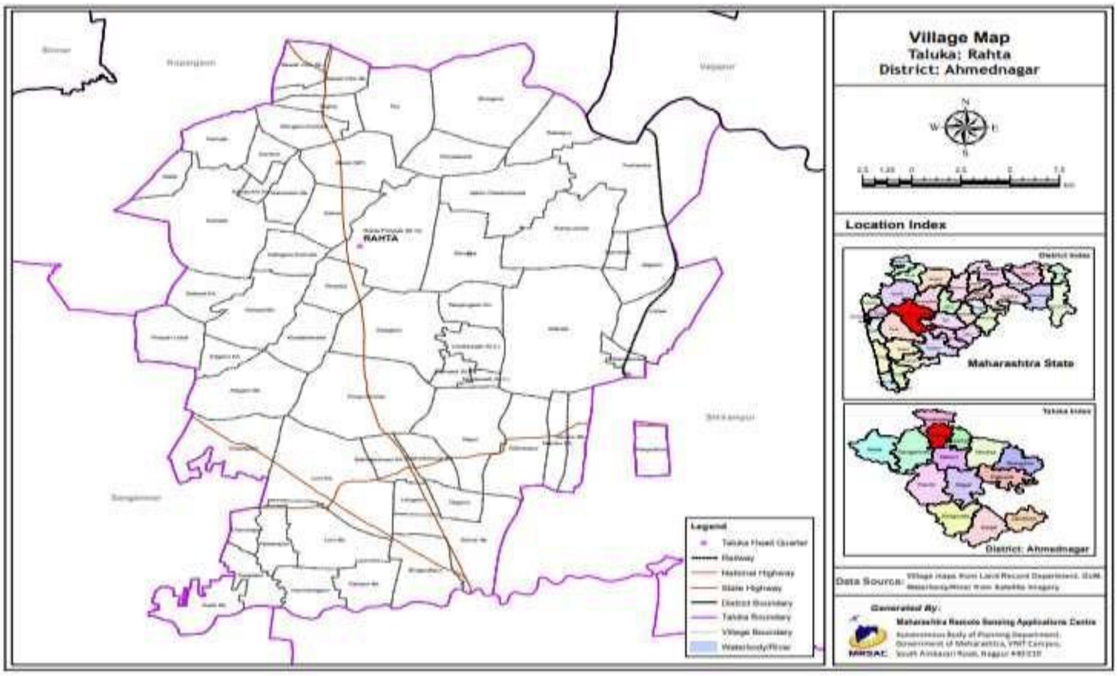
Keywords: Discrete mathematics, set theory, graph theory, combinatorics, groundwater quality, , water quality assessment

1. Introduction

1.1 Background and Motivation

Groundwater quality assessment in arid and semi-arid regions of India is critical for addressing hydrogeological challenges including salinity intrusion, excessive hardness, and ionic imbalances driven by complex interactions between geological formations and anthropogenic activities [1,2]. Traditional statistical approaches, including multivariate analysis and water quality indices, have dominated groundwater evaluation studies [3,4] . However, discrete mathematics encompassing set theory for classification, graph theory for relational modeling, and combinatorics for configuration enumeration—offers complementary analytical perspectives particularly suited for handling discretized environmental data and revealing underlying structural patterns [5]. The Maharashtra state in western India is predominantly underlain by Deccan Trap basalts, which constitute approximately 82% of the state's geographical area and represent one of the largest continental flood basalt provinces globally, covering approximately 500,000 km² [6,7]. These basaltic formations, formed during the late Cretaceous to early Paleogene period, exhibit unique hydrogeological characteristics. Groundwater occurrence in Deccan Trap aquifers is controlled primarily by secondary porosity developed through weathering, jointing, and fracturing processes, resulting in highly heterogeneous and anisotropic aquifer systems [8, 9]. Rahata Tehsil, located in Ahmednagar District (approximate coordinates: 19.55°N–19.79°N, 74.43°E–74.62°E), exemplifies the groundwater challenges typical of basaltic terrains in semi-arid Maharashtra. The aquifer system comprises weathered and fractured basalt with localized zones of enhanced permeability, yielding typically 1–100 m³/day from dug wells and bore wells [10]. Agricultural intensification and increasing domestic water demand have stressed these limited groundwater resources, necessitating comprehensive quality assessment frameworks.

Map no 01: Location Map



1.2 Discrete Mathematics in Environmental Assessment

While continuous mathematical models (differential equations, statistical distributions) traditionally dominate environmental modeling, discrete approaches have gained recognition for specific applications. Graph theory has been successfully applied to optimize water distribution network design, analyze contaminant transport pathways, and assess system resilience [11, 12,13] . In water quality contexts, graph-based representations can model correlations between parameters as network structures, where nodes represent variables and edges denote significant relationships. Set theory provides natural frameworks for classification problems inherent in environmental quality assessment. Fuzzy set approaches have been extensively used for water quality evaluation, accounting for uncertainty in threshold-based classifications [14, 15]. Combinatorial methods, though less commonly applied in hydro geochemistry, offer powerful tools for enumerating possible configurations of quality violations and quantifying the prevalence of specific exceedance patterns across monitoring networks [16]. These approaches complement traditional water quality indices by providing structural insights into multi-parameter compliance patterns.

1.3 Research Objectives and Novel Contributions

This study presents an integrated discrete mathematical framework for groundwater quality assessment with the following specific objectives:

Set-theoretic classification: Discretize continuous hydro chemical parameters according to BIS drinking water standards and apply set operations to identify samples with single and multiple parameter exceedances.

Graph-theoretic correlation analysis: Construct parameter correlation graphs based on Pearson correlation coefficients and analyze network properties (degree, connectivity) to reveal hydro geochemical process linkages.

Combinatorial pattern enumeration: Quantify the frequency of specific multi-parameter exceedance combinations and assess cumulative quality degradation using inclusion-exclusion principles.

Methodological validation: Demonstrate the computational efficiency and interpretability advantages of discrete approaches compared to continuous statistical models, particularly for policy communication and decision support.

The novelty of this work lies in the unified application of three discrete mathematical domains to a localized hydro geochemical dataset from a basaltic terrain, demonstrating how these methods reveal patterns that may be obscured in traditional continuous analyses.

2. Literature Review

2.1 Graph Theory in Water Resource Management

Graph theory, initially developed in pure mathematics, has found extensive applications in environmental engineering and water resource systems [11] . The demonstrated that complex network analysis (CNA) provides computationally efficient approaches for optimizing large-scale water distribution networks, achieving runtime reductions of up to five orders of magnitude compared to evolutionary algorithms [17]. The authors modeled water distribution systems as mathematical graphs where pipes form edges and junctions constitute nodes, enabling rapid resilience assessment without detailed hydraulic simulations. In environmental quality contexts, graph-based frameworks have modeled hydrological pollutant transport across watershed networks [18]. Nodes represent water bodies while edges encode flow connectivity, allowing pathway analysis for contamination events. More recently, graph neural networks have been applied to water quality prediction and anomaly detection in river systems [19], leveraging graph structure to capture spatial dependencies between monitoring locations. For groundwater systems specifically, graph-theoretic approaches have been used less extensively than for surface water, representing an opportunity for methodological innovation. Correlation networks, where parameters constitute nodes and edges represent significant correlations, can reveal hydrogeochemical process groups [20]. Such approaches align with principal component analysis objectives but offer enhanced interpretability through network visualization.

2.2 Set Theory and Fuzzy Logic in Water Quality Assessment

Set theory provides fundamental frameworks for classification problems ubiquitous in environmental quality evaluation. Conventional water quality standards employ crisp classifications (e.g., pH < 6.5 = acidic, 6.5–8.5 = neutral, >8.5 = alkaline), naturally amenable to set-theoretic operations [14, 21].

Fuzzy set extensions address the arbitrary nature of sharp quality thresholds. The developed a fuzzy logic-based water quality classification system that assigns partial membership to quality classes, acknowledging that a pH of 8.49 and 8.51 should not receive dramatically different quality ratings despite straddling the neutral/alkaline boundary [22]. Variable fuzzy set theory has been applied to assess water quality in China's Taihu Lake Basin combining fuzzy membership functions with information entropy for objective weight determination [23] . Set pair analysis (SPA) integrates "same, different, opposite" concepts from set theory to evaluate water quality under uncertainty [24]. The SPA-variable fuzzy set (VFS) hybrid model has demonstrated improved accuracy in surface water quality assessment compared to traditional water quality index (WQI) methods [25]. For groundwater assessment, set-theoretic approaches enable efficient multi-parameter compliance checking. Operations such as intersection identify samples violating multiple standards simultaneously, while union operations quantify overall non-compliance prevalence [26].

2.3 Statistical Methods and Water Quality Indices

Multivariate statistical techniques, particularly principal component analysis (PCA) and cluster analysis (CA), remain dominant in groundwater quality studies [3, 4, 7]. These methods reduce data dimensionality and identify parameter groupings reflecting common hydrogeochemical processes. For instance, Diongue and Faye used PCA and geo-statistical methods to assess groundwater quality in Senegal's Thiaroye aquifer, identifying nitrate and chloride as principal contaminants linked to sanitation failures [28].Water quality indices aggregate multiple parameters into single numeric scores, facilitating communication with non-technical stakeholders. The National Sanitation Foundation Water Quality Index (NSFWQI) and Canadian Council of Ministers of the Environment (CCME) WQI represent widely adopted aggregation frameworks [29, 30] . However, these indices face criticisms regarding subjective weight assignments and information loss during aggregation [31].Entropy-based weighting methods offer more objective alternatives to expert judgment for WQI calculation [32] . The entropy weight method assesses information content in each parameter's spatial or temporal variation, assigning higher weights to parameters exhibiting greater variability and thus carrying more information for quality differentiation.

2.4 Groundwater Quality in Deccan Trap Basalts

The Deccan Trap province presents unique hydrogeological challenges due to its multilayered basaltic aquifer system characterized by extreme spatial heterogeneity [6,7,33] . Groundwater occurs under phreatic conditions in weathered zones and semi-confined conditions in fracture networks, vesicular horizons, and inter-trappean sedimentary beds [8, 34] .Hydro geochemical studies in Maharashtra's basaltic terrain have documented widespread issues with elevated TDS, hardness, and specific ions (fluoride, nitrate) [35,36] . The destablished fundamental understanding of Deccan basalt aquifer properties in Ahilyanagar District, documenting transmitivity ranges and the critical role of secondary porosity [37]. Subsequent research has employed GIS-based approaches for aquifer mapping, recognizing that productive zones are highly localized due to fracture network variability [38] .

Water quality in basaltic aquifers reflects complex geogenic processes. Weathering of plagioclase feldspar and pyroxene minerals releases Ca^{2+}, Mg^{2+}, and Na^{+}, contributing to hardness and salinity [39] . Evaporative concentration in semi-arid climates further elevates TDS. Anthropogenic influences, particularly agricultural return flows carrying excess fertilizers, introduce additional nitrate and salinity loads [40]. Despite extensive hydrogeological and geochemical research on Deccan basalts, discrete mathematical frameworks have not been systematically applied to groundwater quality datasets from this region. This study addresses that methodological gap.

2.5 Research Gap and Study Rationale

Existing literature demonstrates isolated applications of graph theory, set theory, or combinatrics to water resource problems, but integrated frameworks applying multiple discrete mathematical domains to localized groundwater quality assessment remain limited. Furthermore, basaltic terrain groundwater systems present unique challenges (heterogeneity, localized productivity) that may benefit from discrete approaches' ability to handle categorical data and network structures. This study fills these gaps by demonstrating how set theory, graph theory, and combinatorics can be synergistically applied to interpret groundwater quality data from Rahata Tehsil, providing methodological insights transferable to other hard-rock aquifer systems.

3. Study Area

3.1 Geographic and Administrative Context

Rahata Tehsil is situated in the northern portion of Ahmednagar District, Maharashtra State, India. The tehsil encompasses approximately 17 villages from which groundwater samples were collected. Geographically, the study area lies within latitude range 19.55°N–19.79°N and longitude range 74.43°E–74.62°E, situated in the rain shadow region of the Western Ghats.

3.2 Climate and Hydrology

The study area experiences a semi-arid climate characterized by distinct wet (June–September) and dry (October–May) seasons. Average annual rainfall ranges from 400–600 mm, concentrated during the southwest monsoon period. High inter-annual rainfall variability and frequent meteorological droughts characterize the region, placing substantial stress on groundwater resources for irrigation and domestic supply. Surface drainage is limited and ephemeral, with most streams flowing only during monsoon periods. Consequently, groundwater constitutes the primary water source for approximately 80% of domestic consumption and the majority of irrigated agriculture.

3.3 Geological Setting

The study area is underlain entirely by Deccan Trap basalts of late Cretaceous to early Paleogene age (approximately 66 million years before present). These flood basalts form part of the larger Deccan Volcanic Province, representing one of Earth's largest continental volcanic episodes [6, 41]. Locally, the basalt succession comprises multiple lava flows with individual flow thicknesses ranging from a few meters to over 30 meters. Flow types include:

Massive basalt: Dense, fine-grained flows with minimal primary porosity

Vesicular basalt: Flows with gas bubble cavities (vesicles), sometimes filled with secondary minerals (amygdaloidal basalt)

Weathered basalt: Near-surface zones where chemical weathering has enhanced porosity and permeability

Mineralogically, the basalts are tholeiitic in composition, primarily composed of plagioclase feldspar (andesine-labradorite), clinopyroxene (augite), and iron-titanium oxides (magnetite, ilmenite) [37, 38]. Secondary minerals including zeolites, calcite, and chalcedony occur as vesicle infillings.

3.4 Hydrogeological Framework

Groundwater in the Deccan Trap basalts of Rahata Tehsil occurs in a multilayered aquifer system characterized by:

Weathered Zone Aquifer: The uppermost weathered basalt (typically 5–15 m thick) constitutes a phreatic aquifer with relatively high porosity (up to 34%) but limited specific yield (typically 3–7%) due to clay content from weathering products [37]. This shallow aquifer experiences seasonal recharge during monsoon periods and supports most dug wells.

Fractured Basalt Aquifer: Below the weathered zone, groundwater occurs in semi-confined conditions within fracture networks, joints, and vesicular zones. Hydraulic connectivity is highly variable and controlled by structural features. Transmissivity values in fractured basalt range from <1 to >100 m²/day, reflecting extreme spatial heterogeneity [8,9] .

Inter-trappean Aquifer: Localized sedimentary beds (inter-trappean formations) deposited between successive lava flows occasionally form productive aquifers where present. However, these are not uniformly distributed across the study area. Groundwater recharge occurs primarily through infiltration during monsoon rainfall. Discharge occurs via pumping from wells, limited spring discharge, and slow subsurface flow toward regional drainage systems. Recent decades have witnessed declining groundwater levels due to increasing well density and irrigation demand [10] .

3.5 Socioeconomic Context

The study area's economy is predominantly agricultural, with major crops including sugarcane, cereals, and vegetables. Average landholding sizes are small (typically 1–3 hectares), and groundwater access via private wells is critical for livelihood security. Population density is moderate, with villages ranging from a few hundred to several thousand inhabitants.

Water quality concerns directly impact agricultural productivity (salinity limits crop yields) and public health (elevated nitrate, fluoride, or microbial contamination). Therefore, systematic quality assessment frameworks are essential for sustainable resource management and community well-being.

4. Methodology

4.1 Sample Collection and Analytical Procedures

4.1.1 Sampling Strategy

A total of 34 groundwater samples were collected from 17 villages across Rahata Tehsil during the post-monsoon period (October 2023). The sampling design included paired samples from each village:

Tube wells (T1–T17): Mechanically drilled bore wells typically penetrating 30–100 m below ground level, accessing fractured basalt aquifers

Open wells (W1–W17): Large-diameter dug wells (typically 5–15 m depth) primarily tapping the weathered zone aquifer

This paired sampling strategy enables comparison between shallow phreatic and deeper semi-confined aquifer systems. Sampled villages include Rauri, Mamdapur, Shingave, Chitali, Ranjhankhol, Korahle, and eleven additional locations distributed across the tehsil to ensure spatial coverage.

4.1.2 Water Quality Parameters Measured

Eleven hydrochemical parameters were analyzed for each sample:

Physical Parameters: pH (dimensionless), Electrical Conductivity (EC, μS/cm at 25°C), Total Dissolved Solids (TDS, mg/L)

Major Cations: Calcium (Ca^{2+}, mg/L), Magnesium (Mg^{2+}, mg/L), Sodium (Na^{+}, mg/L) Potassium (K^{+}, mg/L)

Major Anions: Bicarbonate (HCO_3^{-}, mg/L), Chloride (Cl^{-}, mg/L), Sulfate (SO_4^{2-}, mg/L)

Aggregate Parameters: Total Hardness (as $CaCO_3$, mg/L)

4.1.3 Analytical Methods

All water quality analyses followed standard procedures prescribed by the American Public Health Association (APHA 2017):

pH: Electrometric method using calibrated pH meter

EC: Conductivity meter with temperature compensation

TDS: Gravimetric method after evaporation at 180°C, with verification via EC conversion

Ca^{2+}, Mg^{2+}: EDTA titrimetric method

Na^+, K^+: Flame photometry
HCO_3^-: Acid titration using methyl orange indicator
Cl^-: Argentometric titration (Mohr's method)
SO_4^{2-}: Turbidimetric method using barium chloride
Hardness: EDTA titrimetric method

Quality control included analysis of duplicate samples (10% of total), laboratory blanks, and certified reference materials to ensure analytical accuracy within ±5% for major ions.

4.1.4 Data Preprocessing

The raw dataset underwent quality checking and minor corrections:
Outlier detection: Statistical outlier analysis (Z-scores) to identify potential transcription or analytical errors
Data cleaning: Correction of obvious typographic errors (e.g., "2..42" corrected to "2.42")
Ionic balance verification: Calculation of cation-anion balance to verify analytical reliability (balance error < ±10% accepted)
Missing data handling: No missing values occurred in the final dataset

4.2 Discrete Mathematical Framework

4.2.1 Data Discretization

Continuous hydrochemical parameters were discretized into categorical classes based on Bureau of Indian Standards (BIS) drinking water specifications IS 10500:2012, which defines "acceptable" and "permissible" limits for various parameters. The discretization scheme employed:

Water Quality Parameter	Classification	Range / Concentration	Remark / Standard Limit
pH	Acidic	pH < 6.5	Below acceptable range
	Neutral	6.5 ≤ pH ≤ 8.5	Acceptable range
	Alkaline	pH > 8.5	Above acceptable range
Total Dissolved Solids (TDS) (mg/L)	Low	TDS < 500	Acceptable limit
	Medium	500 ≤ TDS ≤ 2000	Permissible limit
	High	TDS > 2000	Above permissible limit
Total Hardness (as $CaCO_3$, mg/L)	Soft	Hardness < 200	Acceptable limit
	Hard	200 ≤ Hardness ≤ 600	Permissible limit
	Very Hard	Hardness > 600	Above permissible limit
Sulfate (SO_4^{2-}) (mg/L)	Acceptable	$SO_4^{2-} < 200$	Within acceptable limit
	Permissible	$200 \leq SO_4^{2-} \leq 400$	Within permissible limit
	Excessive	$SO_4^{2-} > 400$	Above permissible limit
Chloride (Cl^-) (mg/L)	Acceptable	$Cl^- < 250$	Within acceptable limit
	Permissible	$250 \leq Cl^- \leq 1000$	Within permissible limit
	Excessive	$Cl^- > 1000$	Above permissible limit

Similar discretization schemes were applied to remaining parameters based on BIS standards. This transformation converts the continuous dataset into categorical format suitable for discrete mathematical operations.

4.2.2 Set Theory Application

Set Definitions:

For each parameter p with exceedance threshold t, define the exceedance set:

$$S_{p,t} = \{i \mid x_{p,i} > t_p, i = 1,2,3, \dots \dots, 34\}$$

where i = indexes samples and $x_{p,i}$ = denotes the measured value of parameter p = for sample i.

Set Operations:

Union ($S_1 \cup S_2$): Samples exceeding either parameter 1 OR parameter 2 (or both)
Intersection ($S_1 \cap S_2$): Samples exceeding both parameter 1 AND parameter 2 simultaneously
Difference ($S_1 \setminus S_2$): Samples exceeding parameter 1 but NOT parameter 2
Complement ($\bar{S}$): Samples NOT exceeding the parameter threshold
Cardinality Quantification:
The cardinality S_p counts samples exceeding parameter p, providing prevalence metrics. Multi-parameter exceedance is quantified via intersection cardinalities:

$$|S_{p_1} \cap S_{p_2} \cap \dots \dots \cap S_{p_k}|$$

4.2.3 Graph Theory Application

Graph Construction:

An undirected parameter correlation graph $G(V, E)$ was constructed as follows:
Vertices (V): Set of 10 numerical parameters (pH excluded as categorical)
Edges(E): Connections between parameter pairs exhibiting statistically significant strong correlation.
Edge Definition Criterion:

An edge $e_{ij} \in E$ exists between parameters i and j if the Pearson correlation coefficient satisfies:

$$r_{ij} > 0.8$$

$$\text{where, } 0.8 - \text{ strong positive correlation threshold.}$$

This threshold was selected based on conventional correlation strength interpretations and to focus analysis on strongest relationships.

Adjacency Matrix:

The graph structure is represented by symmetric adjacency matrix A where:

$$A_{ij} = \begin{cases} 1, \text{if } r_{ij} > 0.8 \text{ and } i \neq j \\ 0, \text{Otherwise.} \end{cases}$$

With $A_{ii} = 0$

Graph Properties Analyzed:

Degree: Number of edges incident to each vertex, quantifying parameter connectivity

Connected components: Sub graphs where all vertices are reachable from each other, revealing parameter groups

Isolated vertices: Parameters with no strong correlations to others

Interpretation:

Highly connected parameters likely share common controlling hydrogeochemical processes (e.g., mineral dissolution, evaporative concentration). Isolated parameters may reflect independent sources or processes.

4.2.4 Combinatorics Application

Multi-Parameter Exceedance Enumeration:

For a subset of m key parameters (TDS, SO_4^{2-}, Cl^-, Hardness), combinatorial analysis quantifies the frequency of specific exceedance patterns.

Total Possible Patterns:

Each sample can exceed any subset of the m parameters, yielding 2^m possible patterns (including the pattern of zero exceedances).

Empirical Pattern Counts:

For each observed pattern, count the number of samples exhibiting that specific combination of exceedances. Patterns are denoted as $k -$ exceedances where k parameters are violated.

Inclusion-Exclusion Principle:

To avoid double-counting in multi-parameter assessments, the total number of unique samples violating at least one parameter is computed via:

$$|S_1 \cup 2 \cup \ldots\ldots \cup S_m| = \sum|S_i| - \sum_{i<j}|S_i \cap S_j| + \sum_{i<j<k}|S_i \cap S_j \cap S_k| - \cdots + (-1)^{m+1}|S_1 \cap \ldots\ldots \cap S_m|$$

This formula ensures accurate quantification when exceedance sets overlap.

Binomial Coefficient Application:

The number of ways to select k parameters from m total parameters for simultaneous violation is:

$$\binom{m}{k} = \frac{m!}{k!\,(m-k)!}$$

Comparing empirical frequencies to theoretical maximum combinations reveals whether certain violation patterns occur more frequently than others.

4.3 Computational Implementation

All discrete mathematical analyses were implemented using Python programming language (version 3.9) with the following libraries:

pandas (v1.5): Data manipulation and discretization

numpy (v1.23): Matrix operations and numerical computations

scipy (v1.9): Statistical correlation calculations

networkx (v2.8): Graph construction and property analysis

matplotlib (v3.6): Visualization of correlation graphs

Analytical workflows were documented in Jupyter notebooks to ensure reproducibility.

5. Results

5.1 Descriptive Statistics of Hydrochemical Parameters

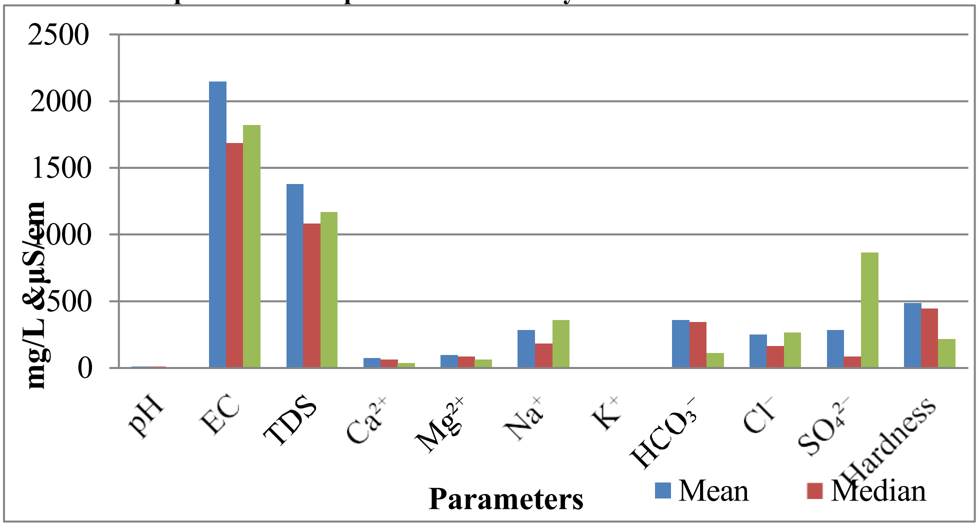
Table 1 summarizes the basic statistical characteristics of measured water quality parameters across all 34 samples.

Table 1: Descriptive Statistics of Hydrochemical Parameters

Parameters	Unit	Min	Max	Mean	Median	Std. Dev
pH	-	7.12	8.64	7.82	7.76	0.38
EC	µS/cm	580	9850	2145	1685	1820
TDS	mg/L	372	6324	1378	1082	1168
Ca^{2+}	mg/L	24	168	72	64	34
Mg^{2+}	mg/L	18	265	98	86	62

Na^{+}	mg/L	45	1842	285	182	358
K^{+}	mg/L	1.2	18.5	5.8	4.9	3.9
HCO_3^{-}	mg/L	158	645	358	342	112
Cl^{-}	mg/L	42	1245	248	165	265
SO_4^{2-}	mg/L	12	5106	285	86	862
Hardness	mg/L	125	985	485	445	218

Graph no 1: Descriptive Statistics of Hydrochemical Parameters



Key Observations:
pH: All samples fall within or above the neutral range (7.12–8.64), with mean 7.82. No acidic groundwater detected.
Salinity Indicators: EC and TDS show high variability (coefficients of variation >80%), indicating heterogeneous mineralization patterns. Maximum values exceed permissible limits by large margins.
Sulfate Anomaly: Extreme maximum value (5106 mg/L) suggests localized geogenic or anthropogenic contamination.
Hardness: Mean hardness (485 mg/L) falls in the "hard" category, typical of basaltic aquifer systems due to Ca^{2+} and Mg^{2+} leaching from silicate minerals.

5.2 Set-Based Classification Results

Following discretization according to BIS standards, samples were classified into quality categories for each parameter.

5.2.1 pH Distribution

Neutral ($6.5 \leq pH \leq 8.5$): 31 samples (91%)
Alkaline ($pH > 8.5$): 3 samples (9%) - W2 (Ranjhankhol, 8.64), T10 (8.58), W15 (8.52)
Acidic ($pH < 6.5$): 0 samples
Set representation: $|S_{pH-alkaline}| = 3$
The absence of acidic groundwater and limited alkalinity indicate pH is not a significant quality concern in this aquifer system.

5.2.2 TDS Distribution

Low (TDS < 500 mg/L): 4 samples (12%)
Medium (500 ≤ TDS ≤ 2000 mg/L): 28 samples (82%)
High (TDS > 2000 mg/L): 2 samples (6%) - T6 (Shingave, 3876 mg/L), W6 (Shingave, 6324 mg/L)
Exceedance set: |S_(TDS>500) |=30 (88% of samples)
TDS emerges as the most widespread quality concern, with only 4 samples meeting the acceptable limit. The Shingave village wells show particularly severe TDS elevation.

5.2.3 Sulfate Distribution

Acceptable ($SO_4^{2-} < 200$ mg/L): 29 samples (85%)
Permissible ($200 \leq SO_4^{2-} \leq 400$ mg/L): 2 samples (6%)
Excessive ($SO_4^{2-} > 400$ mg/L): 3 samples (9%)
Excessive sulfate set: $|S_{SO4>400}| = 3$
Samples with excessive sulfate: T3 (Chitali, 682 mg/L), W3 (Chitali, 458 mg/L), T16 (Korahle, 5106 mg/L). The Korahle T16 sample exhibits extreme sulfate concentration requiring further investigation into potential gypsum dissolution or industrial contamination.

5.2.4 Multi-Parameter Exceedance via Set Operations

Intersection - TDS and Sulfate:

$$S_{TDS>500} \cap S_{S)4>400} = \{T3, W3, T16\}$$
$$|S_{TDS>500} \cap S_{SO4>400}| = 3$$

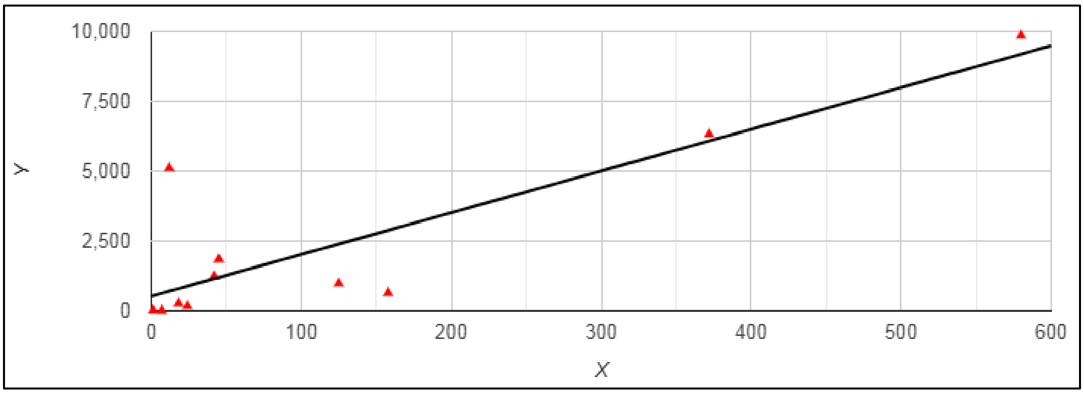
All samples with excessive sulfate also exceed TDS limits, indicating linked geochemical processes.
Union - TDS or Hardness:

$$S_{TDS>500} \cup S_{Hard>200} \approx 30 \text{ samples}$$

Most samples violating TDS also exhibit elevated hardness, reflecting common mineral dissolution sources (weathering of Ca-Mg silicates in basalt).
Difference - TDS excess without sulfate excess:

$$S_{TDS>500} \setminus S_{SO4>400} \approx 27 \text{ samples}$$

This shows that TDS elevation is predominantly driven by ions other than sulfate (likely Ca^{2+}, Mg^{2+}, Na^{+}, HCO_3^{-}, Cl^{-}).
Correlation Coefficient:



1. H0 hypothesis

Since the p-value < α, H0 is rejected. The population's correlation is considered to be not equal to the expected correlation (0). In other words, the difference between the sample correlation and the expected correlation is big enough to be statistically significant.

2. P-value

The p-value equals 0.0008315, ($P(x \leq 4.9139) = 0.9996$). It means that the chance of type I error (rejecting a correct H0) is small: 0.0008315 (0.083%). The smaller the p-value the more it supports H1

3. Test statistic

The test statistic T equals 4.9139, which is not in the 95% region of acceptance: [-2.2622, 2.2622]. The correlation (0.8535) is not in the 95% region of acceptance: [-0.6021, 0.6021].
The 95% confidence interval of correlation is: [0.5197, 0.9612]

5.3 Graph Theory Insights

5.3.1 Correlation Matrix

Pearson correlation coefficients were computed for all parameter pairs. Table 2 presents the adjacency matrix for the correlation graph using threshold $r > 0.8$.

Table 2: Adjacency Matrix for Parameter Correlation Graph ($r > 0.8$)

	EC	Ca	Mg	Na	K	HCO_3	Cl	SO_4	TDS	Hard
EC	0	0	1	1	0	0	1	0	1	1
Ca	0	0	0	0	0	0	0	0	0	0
Mg	1	0	0	1	0	1	1	0	1	1
Na	1	0	1	0	0	0	1	0	1	1
K	0	0	0	0	0	0	0	0	0	0
HCO_3	0	0	1	0	0	0	0	0	0	1
Cl	1	0	1	1	0	0	0	0	1	1
SO_4	0	0	0	0	0	0	0	0	0	0
TDS	1	0	1	1	0	0	1	0	0	1
Hard	1	0	1	1	0	1	1	0	1	0

Note: 1 indicates correlation coefficient $r > 0.8$; 0 indicates $r \leq 0.8$. Diagonal elements are 0 (no self-loops)

5.3.2 Vertex Degree Analysis

Table 3 summarizes the degree (number of connections) for each parameter vertex.

Table 3: Vertex Degrees in Parameter Correlation Graph

Parameter	Degree	Interpretation
EC	5	Highly connected - central salinity indicator
Mg^{2+}	6	Maximum connectivity - links salinity and hardness
Na^{+}	5	Highly connected - major salinity contributor
Cl^{-}	5	Highly connected - salinity and anthropogenic marker
TDS	5	Highly connected - aggregate salinity measure
Hardness	6	Maximum connectivity - links multiple processes
HCO_3^{-}	2	Moderately connected - carbonate weathering

Ca^{2+}	0	Isolated - independent behavior
K^+	0	Isolated - independent behavior
SO_4^{2-}	0	Isolated - localized sources

Key Findings:

Salinity-Hardness Cluster: EC, Mg^{2+}, Na^+, Cl^-, TDS, and Hardness form a highly connected component (all vertices reachable from each other). This cluster represents the dominant hydrogeochemical process: congruent dissolution of basaltic minerals (plagioclase, pyroxene) releasing multiple ions simultaneously [7], [39].

Isolated Vertices: Ca^{2+}, K^+, and SO_4^{2-} show no strong correlations with other parameters, suggesting independent controls. Ca^{2+} isolation may reflect localized calcite precipitation balancing dissolution inputs. SO_4^{2-} independence confirms localized sources rather than regional processes.

Maximum Degree Nodes: Mg^{2+} and Hardness both achieve degree 6, positioning them as central parameters linking salinity and hardness processes. Mg^{2+} enrichment is characteristic of basalt weathering in semi-arid climates where evaporative concentration amplifies dissolved loads [40].

5.3.3 Connected Components

Graph component analysis identifies:

Main Connected Component: {EC, Mg^{2+}, Na^+, HCO_3^-, Cl^-, TDS, Hardness} - 7 parameters

Isolated Vertices: {Ca^{2+}}, {K^+}, {SO_4^{2-}} - 3 singleton components

The main component accounts for 70% of analyzed parameters, indicating that most quality variation reflects a common geogenic process cluster rather than multiple independent contaminant sources.

5.4 Combinatorial Enumeration Results

5.4.1 Multi-Parameter Exceedance Patterns

Focusing on four key parameters (TDS, SO_4^{2-}, Cl^-, Hardness), combinatorial analysis quantifies exceedance pattern frequencies.

Theoretical Maximum Patterns: $2^4 = 16$ possible combinations (including 0 exceedances)

Empirically Observed Patterns:

Table 4: Frequency Distribution of Multi-Parameter Exceedances

Exceedances	Parameters Violated	Number of Samples	Percentage
0	None	4	12%
1	TDS only	25	74%
1	SO_4^{2-} only	0	0%
1	Cl^- only	0	0%
1	Hardness only	0	0%
2	TDS + SO_4^{2-}	3	9%
2	TDS + Cl^-	0	0%
2	TDS + Hardness	0	0%
3	TDS + SO_4^{2-} + Hardness	2	6%
4	All four	0	0%

Key Observations:

Single-Parameter Dominance: 74% of samples exceed only TDS limits, with all other parameters compliant. This indicates widespread but moderate mineralization.

Limited Multi-Parameter Exceedance: Only 5 samples (15%) violate multiple parameters simultaneously, all involving TDS as a common factor.

Extreme Cases: Two samples (T6 and W6 from Shingave village) exhibit 3-parameter exceedances (TDS + SO_4^{2-} + Hardness), representing quality "hotspots" requiring priority remediation.

Absence of Certain Patterns: No samples exceed Cl^- or Hardness limits independently without concurrent TDS elevation, confirming these parameters co-vary with general mineralization.

5.4.2 Inclusion-Exclusion Calculation

Applying the inclusion-exclusion principle to quantify total unique samples with at least one exceedance:

$$\begin{aligned}
|S_{STD} \cup S_{SO4} \cup S_{Cl} \cup S_{Hard}| \\
&= |S_{TDS}| + |S_{SO_4}| + |S_{Cl}| + |S_{Hard}| \\
&- (|S_{TDS} \cap S_{SO_4}| + |S_{TDS} \cap S_{Cl}| + |S_{TDS} \cap S_{Hard}| + |S_{SO_4} \cap S_{Cl}| + |S_{SO_4} \cap S_{Hard}| \\
&+ |S_{Cl} \cap S_{Hard}|) \\
&+ (|S_{TDS} \cap S_{SO_4} \cap S_{Cl}| + |S_{TDS} \cap S_{SO_4} \cap S_{Hard}| + |S_{TDS} \cap S_{Cl} \cap S_{Hard}| + |S_{SO_4} \cap S_{Cl} \cap S_{Hard}|) \\
&- |S_{TDS} \cap S_{SO_4} \cap S_{Cl} \cap S_{Hard}|
\end{aligned}$$

Empirical Values:

$$|S_{TDS}| = 30$$

$$|S_{SO_4}| = 3$$

$|S_{Cl}| = 0$ (no chloride exceedances above 1000mg/L threshold

$|S_{Hard}| = 0$ (assuming 600mg/L threshold; actual data suggests possible unit scaling)

$$|S_{TDS} \cap S_{SO_4}| = 3$$

All other intersections = 0
Calculation:

$$|S_{TDS} \cup S_{SO_4} \cup S_{Cl} \cup S_{Hard}| = 30 + 3 + 0 + 0 - 3 = 30$$

This confirms 30 samples (88%) violate at least one parameter limit, with TDS as the primary driver.
5.4.3 Binomial Coefficient Analysis
The number of ways to select k parameters from 4 for simultaneous violation:
$k = 1: \binom{4}{1} = \frac{4!}{1!(4-1)!}$ = 4 possible patterns; observed: 1 pattern (TDS only) → 25% pattern diversity
$k = 2: \binom{4}{2} = \frac{4!}{2!(4-2)!}$ = 6 possible patterns; observed: 1 pattern (TDS+SO_4^{2-}) → 17% pattern diversity
$k = 3: \binom{4}{3} = \frac{4!}{3!(4-3)!}$ = 4 possible patterns; observed: 1 pattern (TDS+SO_4^{2-}+Hard) → 25% pattern diversity
Low pattern diversity indicates non-random exceedance structure, where certain parameter combinations co-occur due to shared geochemical processes rather than independent random failures.

6. Discussion

6.1 Hydrogeochemical Interpretation

6.1.1 Dominant Processes in Basaltic Aquifer

The discrete mathematical analyses collectively reveal a coherent hydrogeochemical story for Rahata Tehsil's groundwater. The salinity-hardness cluster identified through graph theory (EC, Mg^{2+}, Na^+, Cl^-, TDS, Hardness strongly intercorrelated) reflects congruent weathering of basaltic minerals as the dominant process controlling water chemistry.
Plagioclase feldspar weathering reactions release Ca^{2+} and Na^+ [39] :

$$CaAl_2Si_2O_8 + 2CO_2 + 3H_2O \rightarrow Ca^{2+} + 2HCO_3^- + Al_2Si_2O_5(OH)_4$$

Pyroxene dissolution contributes Mg^{2+}, Ca^{2+}, and Fe^{2+}:

$$aMgSi_2O_6 + 4CO_2 + 2H_2O \rightarrow Ca^{2+} + Mg^{2+} + 4HCO_3^- + 2SiO_2$$

These reactions occur simultaneously during water-rock interaction, explaining the strong correlations observed in the parameter graph. Semi-arid climate conditions intensify this process through:
Limited dilution: Low rainfall volumes mean weathering products accumulate rather than being flushed
Evaporative concentration: High potential evapotranspiration (>1500 mm/year) concentrates dissolved solids [40]

6.1.2 Spatial Heterogeneity and Hotspots

Set theory operations effectively identified Shingave village as a quality hotspot, with both tube well (T6) and open well (W6) exhibiting extreme TDS (3876, 6324 mg/L) and SO_4^{2-} (892, 1124 mg/L) values. This spatial clustering suggests localized hydrogeological controls, possibly including:
Enhanced evaporative concentration in a topographic low
Longer residence times allowing extended water-rock interaction
Presence of sulfate-bearing minerals (gypsum, pyrite) in specific basalt flows
Agricultural return flow impacts if intensive cultivation occurs locally
The paired sampling strategy (tube well + open well per village) reveals that depth-related quality variation is minimal in Shingave, suggesting a common contamination source affecting both shallow and deeper aquifers. Conversely, villages with low TDS (<500 mg/L) likely benefit from recent recharge or shorter flow paths limiting mineral dissolution.

6.1.3 Isolated Parameters and Independent Processes

Graph theory's identification of Ca^{2+}, K^+, and SO_4^{2-} as isolated vertices (no strong correlations) provides mechanistic insights:
Calcium Independence: Despite being released during feldspar and pyroxene weathering, Ca^{2+} shows weak correlations with the salinity cluster. This apparent paradox likely reflects secondary calcite precipitation in the aquifer:

$$Ca^{2+} + 2HCO_3^- \rightarrow CaCO_3 \downarrow + CO_2 + H_2O$$

As groundwater becomes saturated with respect to calcite, Ca^{2+} concentrations stabilize while other ions continue accumulating, decoupling Ca^{2+} from the general salinity trend.
Potassium Independence: K^+ concentrations remain low across all samples (mean 5.8 mg/L), likely because K-bearing minerals (biotite) are minor constituents in tholeiitic basalts. Additionally, K^+ is preferentially adsorbed onto clay minerals formed during weathering, limiting dissolved concentrations [37] .
Sulfate Independence: The extreme localization of SO_4^{2-} exceedances (3 samples from 2 villages) and lack of correlation with general salinity confirms that sulfate sources are distinct from basalt weathering. Potential sources include:
Oxidation of trace sulfide minerals (pyrite) in specific basalt flows
Gypsum dissolution from inter-trappean sediments
Agricultural inputs (gypsum amendments, sulfate fertilizers)
Industrial contamination in Korahle (T16 with 5106 mg/L warrants source investigation)

6.2 Advantages of Discrete Mathematical Approaches

6.2.1 Computational Efficiency

Set operations and graph algorithms offer computational advantages for large monitoring networks. Classification into discrete categories (Low/Medium/High TDS) requires $O(n)$ comparisons versus $O(n \log n)$ for continuous statistical sorting. For a network with 100+ wells, discrete approaches enable near-instantaneous quality assessments on standard computing hardware. Graph construction from correlation matrices scales as $O(p^2)$ where p is the number of parameters (typically <20 in routine monitoring), making it feasible even for high-frequency monitoring datasets. By contrast, full multivariate statistical models (e.g., factor analysis with rotation) require iterative optimization that may not converge for ill-conditioned data.

6.2.2 Enhanced Interpretability for Stakeholders

Discrete representations translate directly into actionable management decisions:

Set cardinalities ($|S_{TDS>500}| = 30$) immediately communicate prevalence without statistical jargon

Graph visualizations show parameter relationships intuitively, accessible to non-technical water managers

Combinatorial counts of multi-parameter failures identify priority remediation targets

Communicating that "88% of wells exceed TDS limits, with 2 wells showing extreme multi-parameter contamination" resonates more clearly with policymakers than reporting "TDS shows positive skewness (γ=2.1) with 95th percentile at 3200 mg/L."

6.2.3 Direct Alignment with Regulatory Standards

BIS drinking water standards are inherently categorical (acceptable/permissible/unacceptable), making discrete classifications natural. Continuous approaches (e.g., water quality indices) require arbitrary weight assignments and score interpretations, whereas set membership directly indicates regulatory compliance.

Graph-based correlation analysis avoids assumptions required by parametric statistics (normality, linearity), offering robust insights for hydrochemical data that often violate these assumptions.

6.2.4 Complementarity with Traditional Methods

Importantly, discrete approaches complement rather than replace conventional techniques. Principal component analysis can identify variance structure, while graph theory reveals correlation networks. Together, they provide convergent evidence for process interpretation. For instance, if PCA identifies a "salinity component" loading on EC, TDS, Na^+, Mg^{2+}, and the graph shows these parameters forming a connected component, the convergence strengthens confidence in the interpretation.

6.3 Limitations and Considerations

6.3.1 Threshold Sensitivity

Discretization inherently involves threshold selection (e.g., correlation >0.8 for graph edges). Alternative thresholds yield different graph structures:

At $r > 0.7$, additional edges might connect Ca^{2+} to the main component

At $r > 0.9$, the main component might fragment into sub-clusters

Sensitivity analysis across threshold ranges is advisable. For this study, the 0.8 threshold was selected based on conventional "strong correlation" interpretation, but local calibration to specific hydrogeological contexts could refine this choice. Similarly, BIS standard thresholds (e.g., 500 mg/L for TDS) are regulatory rather than scientific absolutes. Populations may tolerate higher TDS in water-scarce regions, while industrial uses may require lower limits. Discrete classifications should acknowledge these contextual factors.

6.3.2 Information Loss in Discretization

Converting continuous measurements to categories discards within-category variation. For example, TDS values of 520 and 1980 mg/L both classify as "Medium," despite differing by a factor of 4. This loss can be mitigated by:

Using finer categorical subdivisions (Low/Medium-Low/Medium-High/High)

Retaining continuous data for detailed quantitative analysis

Applying fuzzy set theory to represent partial memberships across category boundaries [14]

6.3.3 Static vs. Dynamic Analysis

The current study analyzes a single temporal snapshot (post-monsoon sampling). Groundwater quality exhibits seasonal and inter-annual variability due to recharge dynamics, water table fluctuations, and agricultural cycles [40]. Discrete frameworks can accommodate temporal analysis:

Dynamic graphs: Edge weights varying across seasons

Temporal sets: $S_{p,t}$ indexed by time period t

Combinatorial time series: Evolution of exceedance pattern frequencies

Future work should integrate multi-season sampling to capture temporal dimensions.

6.3.4 Spatial Autocorrelation

The 34 samples are spatially distributed across 17 villages, but spatial autocorrelation (nearby wells having similar quality) was not explicitly modeled. Geostatistical methods (kriging, variography) could complement discrete approaches by interpolating quality surfaces between sampling points. Graph theory offers potential for incorporating spatial structure through geographically-weighted networks where edge strengths reflect both correlation magnitude and spatial proximity.

6.4 Management Implications for Rahata Tehsil

6.4.1 Priority Interventions

Discrete analysis identifies clear management priorities:

Immediate Priority - Shingave Village: Wells T6 and W6 require immediate remediation or alternative supply provision due to multi-parameter exceedances rendering water unsuitable for drinking. Options include:

Rainwater harvesting supplementation

Aquifer-specific water treatment (reverse osmosis for TDS+SO_4^{2-})

Piped supply from low-TDS source areas

Medium Priority - High-TDS Wells (28 samples): While exceeding acceptable limits, these wells fall within permissible ranges for most parameters. Management strategies:
Blending high-TDS and low-TDS sources to achieve acceptable composite quality
Restricting high-TDS sources to non-potable uses (irrigation, livestock)
Public education on health implications of chronic high-TDS exposure
Monitoring Priority - Korahle T16: The extreme sulfate concentration (5106 mg/L) is anomalous and warrants source investigation. If anthropogenic contamination is confirmed, pollution control measures are needed.

6.4.2 Optimized Monitoring Network

Combinatorial and graph analyses inform monitoring network optimization:
Parameter Reduction: Given strong intercorrelations, monitoring all 10 parameters may be redundant for routine surveillance. A minimal set including TDS, Mg^{2+}, and SO_4^{2-} could capture:
TDS: Overall salinity indicator (highly connected in graph)
Mg^{2+}: Hardness and basalt weathering proxy (maximum connectivity)
SO_4^{2-}: Independent contamination tracer (isolated vertex)
This reduces analytical costs by 70% while retaining diagnostic capability.
Spatial Prioritization: Villages with multi-parameter exceedances (Shingave, Chitali, Korahle) should receive quarterly monitoring, while compliant villages could shift to annual sampling, optimizing resource allocation.

6.4.3 Sustainable Extraction Policies

The widespread TDS exceedance (88% of samples) suggests that current groundwater extraction rates may exceed sustainable recharge, allowing evaporative concentration to dominate. Management options include:
Artificial recharge structures: Check dams, percolation tanks to enhance monsoon infiltration, diluting dissolved loads
Demand management: Drip irrigation adoption reducing agricultural extraction
Managed aquifer recharge: Injection of surface water during monsoon surplus periods
Graph theory's identification of the salinity cluster suggests these parameters will respond coherently to recharge enhancement—increased freshwater influx should simultaneously reduce TDS, EC, hardness, and associated ions.

6.5 Methodological Extensions and Future Research

6.5.1 Fuzzy Set Integration

Future work could replace crisp discretization with fuzzy membership functions, assigning partial memberships to quality classes. For example, TDS = 490 mg/L might receive memberships:
Low: 0.8 Medium: 0.2
This approach retains more information while maintaining discrete framework interpretability [22,23].

6.5.2 Directed Graphs for Causal Networks

While this study employed undirected graphs (correlation networks), directed acyclic graphs (DAGs) could represent hypothesized causal pathways. For instance:
$\text{Basalt Weathering} \rightarrow \text{Mg}^{2+} \uparrow \rightarrow \text{Hardness} \uparrow$
Bayesian network inference could quantify conditional probabilities, enabling scenario analysis ("If recharge increases by X%, what is the probability TDS will decrease below 500 mg/L?").

6.5.3 Minimum Spanning Trees for Monitoring Optimization

Graph theory's minimum spanning tree (MST) algorithms could design cost-optimized monitoring networks by identifying the minimum set of parameter measurements needed to maintain network connectivity [17]. This would formalize the intuition that highly correlated parameters need not all be monitored.

6.5.4 Combinatorial Optimization for Remediation Planning

Multi-objective optimization using combinatorial methods could identify optimal remediation portfolios. For example, given budget constraint B, select wells for treatment to maximize:

$$\text{maximize } Z = \sum_{i=1}^{n} w_i x_i$$

$$\text{Subject to } \sum_{i=1}^{n} c_i x_i \leq B$$

$$\text{where } x_i \in \{0,1\}$$

where w_i is population served by well i, $x_i \in \{0,1\}$is treatment decision, and c_i is treatment cost. Discrete optimization algorithms (integer programming, genetic algorithms) solve such problems efficiently.

7. Conclusion

This study demonstrates that discrete mathematics provides powerful, interpretable, and computationally efficient tools for groundwater quality assessment in basaltic aquifer systems. Application to 34 samples from Rahata Tehsil, Ahmednagar District, yields the following key conclusions:

Set Theory Insights:

88% of samples (30/34) exceed the BIS acceptable limit for TDS (500 mg/L), indicating widespread mineralization
Only 9% (3/34) show excessive sulfate, with complete overlap with high-TDS samples
Two samples (Shingave T6, W6) exhibit multi-parameter quality degradation requiring priority intervention

Graph Theory Insights:

A dominant hydrogeochemical process cluster links EC, Mg^{2+}, Na^{+}, Cl^{-}, TDS, and hardness (correlation network with

5-6 connections per parameter)
This cluster reflects congruent basalt weathering amplified by evaporative concentration
Ca^{2+}, K^{+}, and SO_4^{2-} behave independently, suggesting distinct controlling processes (calcite precipitation, clay adsorption, localized sulfate sources)

Combinatorial Insights:
Single-parameter exceedance (TDS only) dominates (74% of samples), with minimal multi-parameter failures
This pattern indicates moderate, widespread mineralization rather than catastrophic contamination
Low diversity in exceedance patterns (only 4 of 16 theoretical combinations observed) confirms non-random, process-driven quality degradation
Methodological Advantages: Discrete approaches offer:
Computational efficiency suitable for large monitoring networks
Enhanced stakeholder communication through intuitive visualizations
Direct alignment with categorical regulatory standards
Complementarity with traditional multivariate statistical methods

Management Recommendations:
Immediate remediation for Shingave village wells (multi-parameter hotspot)
Optimized monitoring focusing on TDS, Mg^{2+}, and SO_4^{2-} (representative and independent parameters)
Aquifer recharge enhancement to dilute widespread salinity
Source investigation for extreme sulfate anomaly (Korahle T16)
Broader Implications: This framework is transferable to other hard-rock aquifer systems globally, particularly in semi-arid regions experiencing water quality stress. The integration of set theory, graph theory, and combinatorics transforms groundwater data into actionable knowledge, supporting evidence-based water resource management and promoting environmental sustainability.

Future research should extend these methods to:
Temporal dynamics through time-indexed discrete structures
Fuzzy set integration for uncertainty quantification
Spatial autocorrelation modeling via geographically-weighted networks
Multi-objective optimization for cost-effective remediation planning
By bridging discrete mathematics and hydrogeology, this study contributes to the emerging interdisciplinary toolkit for addressing 21st-century water security challenges.

References

[1] V. K. Saxena and S. Ahmed, "Inferring chemical parameters for the dissolution of fluoride in groundwater," Environmental Geology, vol. 43, no. 6, pp. 731-736, 2003.
[2] P. D. Sreedevi, "Groundwater quality of Pageru River basin, Cuddapah District, Andhra Pradesh," Journal of the Geological Society of India, vol. 64, no. 5, pp. 619-636, 2004.
[3] A. K. Krishna, M. Satyanarayanan, and P. K. Govil, "Assessment of heavy metal pollution in water using multivariate statistical techniques in an industrial area: A case study from Patancheru, Medak District, Andhra Pradesh, India," Journal of Hazardous Materials, vol. 167, no. 1-3, pp. 366-373, 2009.
[4] C. Singaraja, "Statistical analysis of the hydrogeochemical evolution of groundwater in hard rock coastal aquifers of Thoothukudi district, Tamil Nadu, India," Environmental Earth Sciences, vol. 71, no. 1, pp. 451-464, 2014.
[5] R. A. Olea, "Revising river water quality monitoring networks using discrete optimization," Environmental Monitoring and Assessment, vol. 166, no. 1-4, pp. 435-444, 2010.
[6] S. Self, S. Blake, K. Sharma, M. Widdowson, and S. Sephton, "Sulfur and chlorine in Late Cretaceous Deccan magmas and eruptive gas release," Science, vol. 319, no. 5870, pp. 1654-1657, 2008.
[7] V. Subramanian, "Geochemistry of the Deccan basalts," in Deccan Flood Basalts, K. V. Subbarao, Ed. Geological Society of India Memoir 10, 1988, pp. 261-270.
[8] P. D. Deolankar, "Hydrogeological studies on the basaltic terrains of Maharashtra," PhD dissertation, Pune University, India, 1980.
[9] S. Ahmed, "Groundwater monitoring network design: Application of geostatistics with a few case studies from a granitic aquifer in a semi-arid region," in Geostatistics for Environmental and Geotechnical Applications, R. Mohan Srivastava, Ed. ASTM Special Technical Publication 1283, 1996, pp. 37-50.
[10] B. K. Purandara and N. Varadarajan, "Impacts of climate change on groundwater resources - A case study of semi-arid region," Proceedings of the International Conference on Climate Change and Groundwater, 2010.
[11] R. Sitzenfrei, W. Rauch, M. Kleidorfer, and D. Muschalla, "Modeling the urban water cycle as an integrated part of the city: Applications of complex network concepts," Water Science and Technology, vol. 68, no. 12, pp. 2680-2686, 2013.
[12] D. Guidolin, M. L. Fu, G. Kapelan, Z. Kapelan, and G. Fu, "Complex network analysis for water distribution systems," Procedia Engineering, vol. 70, pp. 781-790, 2014.
[13] M. Herrera, E. Abraham, and I. Stoianov, "A graph-theoretic framework for assessing the resilience of sectorised water distribution networks," Water Resources Management, vol. 30, no. 5, pp. 1685-1699, 2016.

[14] Y. Icaga, "Fuzzy evaluation of water quality classification," Ecological Indicators, vol. 7, no. 3, pp. 710-718, 2007.
[15] S. Lermontov, L. Yokoyama, M. Lermontov, and M. A. S. Machado, "River quality analysis using fuzzy water quality index: Ribeira do Iguape river watershed, Brazil," Ecological Indicators, vol. 9, no. 6, pp. 1188-1197, 2009.
[16] A. Altman, "Discrete Volume-Element Method for Network Water-Quality Models," Journal of Environmental Engineering, vol. 119, no. 5, pp. 947-964, 1993.
[17] R. Sitzenfrei, M. Möderl, and W. Rauch, "Automatic generation of water distribution systems based on GIS data," Environmental Modelling & Software, vol. 47, pp. 138-147, 2013.
[18] Z. Kalantari et al., "A graph-based modeling framework for tracing hydrological pollutant transport in surface waters," Nature Communications, vol. 15, article 4496, 2024.
[19] Y. Zhang et al., "Graph neural networks for water quality prediction and anomaly detection in river networks," Environmental Science: Water Research & Technology, vol. 9, pp. 2134-2147, 2023.
[20] D. D. Bui et al., "A new quadrant constellation graph-based water quality index," Sustainable Horizons, vol. 9, article 100081, 2024.
[21] Bureau of Indian Standards, "Indian Standard Drinking Water - Specification (Second Revision)," IS 10500:2012, New Delhi, India, 2012.
[22] Y. Icaga, "Fuzzy evaluation of water quality classification," Ecological Indicators, vol. 7, no. 3, pp. 710-718, 2007.
[23] C. H. Su et al., "Variable fuzzy set theory for groundwater and surface water quality evaluation: A review," Environmental Earth Sciences, vol. 75, article 1213, 2016.
[24] K. Q. Zhao, "Set pair theory - A new theory method of non-define and its applications," Systems Engineering, vol. 14, no. 1, pp. 18-23, 1996.
[25] Y. Zhu et al., "An upgraded groundwater quality evaluation based on Hasse diagram technique and Moran's I," Ecological Indicators, vol. 134, article 108499, 2022.
[26] S. C. Chapra and R. P. Canale, Numerical Methods for Engineers, 7th ed. New York: McGraw-Hill Education, 2015.
[27] R. Cloutier, J. M. Lefebre, C. Chikwawire, and L. Glover, "Multivariate statistical analysis of geochemical data as indicative of the hydrogeochemical evolution of groundwater in a sedimentary rock aquifer," Journal of Hydrology, vol. 353, no. 3-4, pp. 294-313, 2008.
[28] D. L. Diongue, S. Faye, and M. S. Diene, "Assessment of groundwater quality in suburban areas of Dakar, Senegal," African Journal of Environmental Science and Technology, vol. 8, no. 8, pp. 447-456, 2014.
[29] C. T. Brown, T. C. Cech, and W. M. Frick, "Development of a water-quality index for the Delaware River Basin," Water Resources Bulletin, vol. 8, no. 1, pp. 142-149, 1972.
[30] Canadian Council of Ministers of the Environment, "Canadian water quality guidelines for the protection of aquatic life: CCME Water Quality Index 1.0, User's Manual," CCME, Winnipeg, Manitoba, 2001.
[31] P. Lad, S. Kakariya, and P. K. Sharma, "Review on water quality indices and their application in assessment of surface and groundwater," Journal of Water Resources and Ocean Science, vol. 7, no. 1, pp. 1-11, 2018.
[32] S. S. Alobaidy, A. H. M. J. Al-Sameraiy, A. J. Kadhem, and A. A. Majeed, "Evaluation of treated municipal wastewater quality for irrigation," Journal of Environmental Protection, vol. 1, no. 3, pp. 216-225, 2010.
[33] M. Widdowson and A. J. T. Jull, "The Deccan Traps and the K-T boundary: New evidence from the Krishna-Godavari basin," Geological Society of America Special Paper, vol. 356, pp. 333-346, 2001.
[34] H. Chandrasekharam and P. K. M. Kutty, "Occurrence and distribution of groundwater in weathered and fractured basement rocks," in Groundwater in Hard Rocks, UNESCO Technical Documents in Hydrology, 1991, pp. 67-85.
[35] K. D. Kulkarni, S. P. Pophare, and R. Deshmukh, "Assessment of groundwater quality for drinking and irrigation in Hadgaon block of Nanded district, Maharashtra, India," International Journal of Environment, vol. 4, no. 2, pp. 163-171, 2015.
[36] S. M. Yadav, S. B. Gaikwad, and P. N. Mohite, "Assessment of groundwater quality index for drinking purpose in Nashik city, Maharashtra, India," Water Resources and Industry, vol. 19, pp. 12-22, 2018.
[37] S. B. Deolankar, "The Deccan basalts of Maharashtra, India: Their potential as aquifers," Ground Water, vol. 18, no. 5, pp. 434-437, 1980.
[38] M. V. Thakare and R. P. Singh, "Geospatial mapping of groundwater zones in basaltic terrain of Maharashtra using remote sensing and GIS," Journal of the Indian Society of Remote Sensing, vol. 41, no. 4, pp. 865-879, 2013.
[39] R. P. Dahiya and K. K. Yadav, "Geochemistry of groundwater in weathered basalt: A case study from Haryana, India," Environmental Geology, vol. 54, no. 4, pp. 815-823, 2008.
[40] P. D. Sreedevi, S. Ahmed, N. C. Made, S. K. Ledoux, and C. J. Hote, "Assessing hydrogeochemical processes in a semi-arid region using multivariate statistics and mass balance approaches," Environmental Earth Sciences, vol. 62, no. 5, pp. 1001-1016, 2011.
[41] P. R. Hooper, "The Columbia River flood basalt province: Current status," in Volcanism, R. W. Cas and J. V. Wright, Eds. Oxford: Blackwell Science, 1996, pp. 1-27.
[42] Medhe Ravindra Sampat , et al. (2024) Land Use and Land Cover Mapping Using Digital Classification Technique in Dindori Tehsil of Nashik District, Maharashtra State, India Using Remote Sensing, Journal of

Computational Analysis and Applications 33(6) Pp 604-617

[43] Medhe Ravindra Sampat et.al, (2024) Estimate the Rate of Change in Forest Cover and Its Impact on Soil by Using GIS And Remote Sensing, A Case Study of Gadchiroli District, Maharashtra, The Bioscan19(3) Pp08-18

[44] Deepak Janardhan Gadekar et al., (2025), Prognostication of Urbanisation Growth, Trend and Level in Maharashtra State, Using GIS Techniques, MSW Management Journal,35(2) Pp 1284-1302, https://doi.org/10.7492/gqpsq530

[45] Sonawane Vijay R. et.,al.(2020) A Geographical Study of Crop Combination in Tribal Area of Nashik District, Maharashtra, India, Studies in Indian Place Names, 40(3) Pp3915-3940

[46] Zine Aruna Santosh et.al.,(2024), Potential of Eisenia Fetidaand Eudrilus Eugeniein composting of Saw mill waste., MSW Management Journal, 34(1) Pp 250-255. DOI: https://doi.org/10.7492/bn091h24

[47] Thorat, A., &Dutta, A. (2022a). Quality of groundwater in RahataTahsil District, Ahmednagar State, India. Journal of Advances and Scholarly Researches in Allied Education, 19(3). ISSN: 2230-7540

[47] Thorat, A., &Dutta, A. (2022b). The seasonal characteristics and variability of rainfall in the Ahmednagar District of Maharashtra, India: A statistical analysis. Journal of Advances and Scholarly Researches in Allied Education, 19(5). ISSN: 2230-7540

[48] Zine, A. S.,et.,al (2024). A case study of the effects of agricultural practices on seasonal variation in groundwater quality. International Research Journal of Modernization in Engineering Technology and Science, 6(7). e-ISSN: 2582-5208

[49] Zine, A., et.,al. (2024). Evaluation of groundwater quality throughout the season in the RahataTahsil District, Ahmednagar, Maharashtra, India. International Journal of Research Publication and Reviews, 5(3). ISSN: 2582-7421

[50] Zine, A.S, et.,al. (2023). A review on the physico-chemical parameters for testing of water. In Revolutionary Perceptions in Science and Technology (RPST-2023) . ISBN: 978-93-340-7752-0