CHANGE IN SOIL AND GROUNDWATER BEHAVIOUR DUE TO CONTAMINATION FROM LANDFILL LEACHATE

A Thesis Submitted
In Partial Fulfillment of the Requirements
for the Award of the Degree of

DOCTOR OF PHILOSOPHY

in

CIVIL ENGINEERING

by

SHIVANGI BHARADWAJ

(2K21/PHD/CE/03)

Under the Supervision of

PROF. ASHOK KUMAR GUPTA **PROF. ANIL KUMAR SAHU**



Department of Civil Engineering

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-110042, India

October, 2026

©DELHI TECHNOLOGICAL UNIVERSITY, DELHI-2026
ALL RIGHTS RESERVE

I begin this acknowledgement by offering my heartfelt pranam to Lord Shiva, Maa Saraswati, Shri Ganesh, and Shri Krishna, whose divine blessings, wisdom, strength, and guidance have been a constant source of support throughout my PhD journey. I seek the grace of Lord Shiva for inner strength and perseverance, Maa Saraswati for knowledge and clarity of thought, Shri Ganesh for removing obstacles from my path, and Shri Krishna for guiding me towards dedication, patience, and faith in my karma. The teaching of the Bhagavad Gita, especially the principle of performing one's duty sincerely without attachment to the results, has been a constant source of inspiration during my doctoral research. It reminded me to remain focused on effort, discipline, learning, and honesty, especially during challenging phases of research work.

कर्मण्येवाधिकारस्ते मा फलेषु कदाचन।

मा कर्मफलहेतुर्भूर्मा ते सङ्गोऽस्त्वकर्मणि॥

(श्रीमद्भगवद्गीता, द्वितीय अध्याय, श्लोक 47)

I express my sincere and profound gratitude to my PhD supervisors, **Prof. Ashok Kumar Gupta** and **Prof. Anil Kumar Sahu**, for their continuous guidance, valuable suggestions, encouragement, and support throughout the course of my research work. Their academic expertise, constructive feedback, patience, and constant motivation have played a significant role in shaping this research work and bringing it to successful completion. I am deeply grateful for the time, knowledge, and valuable guidance they provided at every stage of my doctoral journey.

I would also like to express my heartfelt thanks to the Head of the Department, **Prof. Awadhesh Kumar**, and the Chairman of the Department Research Committee (DRC), **Prof. Raju Sarkar**, for their valuable support and encouragement during my research work. I am also sincerely grateful to **Prof. K. C. Tiwari**, **Prof. V. K. Minocha**, **Prof. Shilpa Pal**, **Prof. Amit Kr.**

Shrivastava, Prof. A. Trivedi, **Prof. S. Anbu Kumar, Prof. T. Vijay Kumar** and all the faculty members of the department for their cooperation, guidance, and support throughout my research journey. I would also like to extend my special thanks to **Dr. Ritu Raj** for his constant support, encouragement, and motivation during the difficult phases of this journey. I am also sincerely thankful to **Mr. Hrishikesh Dubey** for his kind support and cooperation.

I would also like to express my heartfelt thanks to **Mr. Sunil Tirkey**, Lab Technician, and **Mr. Atul Sharma**, Laboratory Attendant, for their cooperation, technical assistance, and support during the laboratory testing and experimental work carried out as part of my research.

I am thankful to my friends and fellow research scholars, **Mr. Rishabh Tyagi, Dr. Deepak Sharma, Mr Sandeep Baghel, Mr Ashwani Sharma**, and **Ms. Preeti Pateriya**, for their encouragement, valuable discussions, and moral support throughout this journey. Their presence, cooperation, and constant support made the difficult phases of my research more manageable and meaningful.

I owe my deepest gratitude to my family for their unconditional love, blessings, patience, and constant support throughout this long and demanding journey. I am especially thankful to my mother, **Dr. Meera Sharma**, and my mother-in-law, **Mrs. Shashi Tiwari**, for their inspiration, encouragement, and words of strength, which helped me remain determined and focused. I am deeply grateful to my husband, **Mr. Shashank Tiwari**, for his understanding, patience, and unwavering support during every phase of my doctoral journey. I also express my heartfelt love and gratitude to my beloved son, **Karmanya Tiwari**, who, at the tender age of 3.5 years, supported me in his own innocent and loving way. The love, faith, sacrifices, and emotional support of my family have been my greatest strength and motivation in completing this research.

(Shivangi Bharadwaj)

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-42

CANDIDATE'S DECLARATION

I **SHIVANGI BHARADWAJ** hereby certify that the work which is being presented in the thesis entitled **"Change in Soil and Groundwater Behaviour due to Contamination from Landfill Leachate"** in partial fulfillment of the requirements for the award of the Degree of **Doctor of Philosophy**, submitted in the Department of **Civil Engineering**, Delhi Technological University is an authentic record of my own work carried out under the supervision of **Prof. Ashok Kumar Gupta (DTU) and Prof. Anil Kumar Sahu (DTU).**

The matter presented in the thesis has not been submitted by me for the award of any other degree of this or any other Institute.

Candidate's Signature

This is to certify that the student has incorporated all the corrections suggested by the examiners in the thesis and the statement made by the candidate is correct to the best of our knowledge.

Signature of Supervisors **Signature of External Examiner**

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-42

CERTIFICATE BY THE SUPERVISOR(s)

Certified that **SHIVANGI BHARADWAJ** (enrollment no. 2K21/PHDCE/03) has carried out her research work presented in this thesis entitled **"Change in Soil and Groundwater Behaviour due to Contamination from Landfill Leachate"** for the award of **Doctor of Philosophy** from the Department of Civil Engineering, Delhi Technological University, Delhi, under our supervision. The thesis embodies results of original work, and studies are carried out by the student herself, and the contents of the thesis do not form the basis for the award of any other degree to the candidate or to anybody else from this or any other University/Institution.

(Prof. Ashok Kumar Gupta)
Department of Civil Engineering
Delhi Technological University,
Delhi, India

(Prof. Anil Kumar Sahu)
Department of Civil Engineering
Delhi Technological University,
Delhi, India

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-42

PLAGIARISM VERIFICATION

Title of the Thesis	Change in Soil and Groundwater Behaviour due to Contamination from Landfill Leachate
Total Pages	181
Name of the Scholar	Shivangi Bharadwaj
Supervisor	Prof. Ashok Kumar Gupta & Prof. Anil Kumar Sahu
Department	Civil Engineering

This is to report that the above thesis was scanned for similarity detection. Process and outcome is given below:

Software used: Turnitin
Similarity Index: 9 %
Total Word Count: 42009

Date: 17-10-2026

Candidate's Signature **Signature of Supervisors**

ABSTRACT

India is facing a growing challenge of municipal solid waste (MSW) management, with landfilling and open dumping still widely practiced in many urban areas. Delhi, the capital city of India, depends heavily on three major landfill sites, namely Bhalswa, Okhla and Ghazipur, for municipal solid waste disposal. The present study focuses on Bhalswa landfill, an unlined landfill site located in Model Town tehsil of North District, Delhi. The landfill covers an area of about 52 acres (21.06 ha), receives approximately 2500 MT/day of solid waste and has been operational since 1994. It has reached a height of about 60–65 m above the surrounding ground level, raising serious concerns regarding leachate generation, contaminant migration, soil quality deterioration and groundwater degradation. Since the landfill is unlined, leachate may migrate through the subsurface and affect the surrounding soil and groundwater systems. In this context, the present study evaluates the impact of Bhalswa landfill on the surrounding soil and groundwater environment through an integrated assessment of leachate quality, soil chemistry, heavy metal contamination, geotechnical behaviour, groundwater quality, water quality indexing and machine learning-based prediction approaches.

Leachate samples were analyzed to determine their physicochemical and heavy metal characteristics and showed high pollution potential. Soil samples collected from different borehole locations and depths were assessed for physiochemical, heavy metals, geotechnical and microstructural changes. In soil samples, heavy metal enrichment was observed, with Zn, Pb, Fe, Cd and Mn exceeding Indian natural background soil limits, while Cd also exceeded the WHO soil limit. The overall contamination pattern for soil samples followed the order: Zn > Pb > Mn > Cd > Fe > Cu > Cr > Ni > As > Co. Pollution indices further indicated metal contamination and anthropogenic enrichment, with Igeo identifying Cd, Zn and Pb as major metals of concern and

EF showing enrichment of Zn, Pb, Mn and Cd. SEM–EDX analysis provided supporting elemental evidence of contamination within the soil matrix.

The geotechnical assessment showed that, although soil chemistry was affected by leachate interaction, severe or uniform deterioration in geotechnical behaviour was not observed across all boreholes. The soil profile remained predominantly sandy, with noticeable changes mainly restricted to localized fine-rich or deeper layers, where variations in moisture requirement, dry density, permeability and shear strength were recorded.

Groundwater quality assessment in the landfill-adjacent area indicated significant degradation, suggesting possible influence of leachate migration on subsurface water resources. The heavy metal distribution pattern and water quality assessment showed that most groundwater samples were unsuitable for direct use without treatment, including drinking, irrigation and industrial purposes.

Machine learning models were developed for groundwater quality prediction using hydro-chemical parameters. This study employed six ML models, Artificial Neural Network (ANN), Random Forest (RF), K-Nearest Neighbors (KNN), Decision Tree (DT), XGBoost, and Support Vector Machine (SVM) to predict water quality. Among the six models, ANN showed the best performance for complete feature-based water quality prediction, while SVM demonstrated reliable prediction using reduced input features. Validation with independent datasets has confirmed the robustness and applicability of the suggested methodology. The outcomes of the study indicate that ANN is the best performing model, providing a cost-efficient alternative for predicting groundwater quality, significantly reducing extensive laboratory testing and offering practical support for real-time monitoring and helps policy makers to take decisions for mitigation

and improving water quality. Additionally, this study has offered practical implications for scalable, cost-effective deployment in areas with water scarcity or insufficient dataset.

Due to the multidimensionality of the soil dataset, traditional predictive methods proved insufficient in terms of scalability and robustness. To address these gaps, A hybrid machine learning model was developed using Random Forest and Genetic Algorithm (RF-GA). The key findings indicate that the RF-GA model was benchmarked against five established algorithms such as RF, XGBoost, Support Vector Machines (SVM), Artificial Neural Networks (ANN), and Logistic Regression (LR).

Overall, the study provides an integrated framework for assessing landfill-induced soil and groundwater contamination. The findings highlight the need for regular monitoring, leachate control, groundwater treatment and sustainable landfill management to reduce environmental and public health risks in landfill-affected regions.

Keywords Anthropogenic pollution; Soil–environment interactions; Subsurface transport; Soil profile dynamics; Aquifer Recharge; Water Quality; Contaminant Transport; Groundwater Monitoring; Landfill leachate; Artificial Neural Network (ANN), RF-GA hybrid model;

LIST OF PUBLICATIONS

Publications resulting from this thesis work

- Bharadwaj, S., Gupta, A. K., & Sahu, A. K. (2025). Tracing Groundwater Contaminants for Sustainable Development Through Machine Learning Algorithm. *CLEAN – Soil, Air, Water,* 53(12). Wiley. https://doi.org/10.1002/clen.70083 SCIE.
- Bharadwaj, S., Gupta, A. K., & Sahu, A. K. (2025). Innovations in water quality management using machine learning approaches. *Indian Journal of Engineering and Materials Sciences (IJEMS),* 32(05), 536-551. https://doi.org/10.56042/ijems.v32i05.20867 SCIE.
- Bharadwaj, S., Gupta, A. K., & Sahu, A. K. (2026). Optimized Machine Learning model for Predicting Heavy Metal Contamination in Soil around a Landfill. *Zeitschrift der Deutschen Gesellschaft für Geowissenschaften - Journal of Applied and Regional Geology* https://doi.org/10.1127/zdgg/0529 SCIE.
- Jyotirmay, Yadav, A., Meena, J., Bharadwaj, S., & Gupta, A. K., Enhancement of Landfill Leachate Attenuation in Sandy Soil Using Compost Amendment: A Soil Column Study. *CLEAN – Soil, Air, Water,* 53(12). Wiley. https://doi.org/10.1002/clen.70233 SCIE.
- Bharadwaj, S., Gupta, A. K., & Sahu, A. K. Landfill Leachate Induced Groundwater and Surface Water Contamination in Urban Environment. *Sadhana* (under review) SCIE.
- Bharadwaj, S., Gupta, A. K., & Sahu, A. K. Horizontal and Vertical Analysis of Leachate Induced Soil Contamination in and Around a Landfill *Environmental Monitoring and Assessment* (under review) SCIE
- Bharadwaj, S., Gupta, A. K., & Sahu, A. K. (2025). Effect of Leachate Contamination on Soil. Advances in Construction Management, *Lecture Notes in Civil Engineering, 601*

(98–108). Springer. https://doi.org/10.1007/978-981-96-4902-0_9. Scopus.

- Bharadwaj, S., Gupta, A. K., & Sahu, A. K. (2025). Characterization of Contaminated Soil and Groundwater Due to Landfill Leachate. Proceedings of the Indian Geotechnical Conference (IGC 2024), *Lecture Notes in Civil Engineering, 700(91–107).* Springer. https://doi.org/10.1007/978-981-96-7771-9_7 Scopus.

Paper presented in International Scopus indexed conferences

- Bharadwaj, S., Gupta, A. K., & Sahu, A. K. (2023) "Analysis of Landfill Slope Stability in Delhi" In 4th International Conference on waste management, Recycle 2023 at IIT Guwahati in 18-19 May 2023.
- Bharadwaj, S., Gupta, A. K., & Sahu, A. K. (2023) "Effect of Leachate Contamination on Soil" In 7th International Conference on Construction, Real Estate, Infrastructure & Project Management (ICCRIP 2023) At NICMAR Pune, India in 11-12 Aug 2023.
- Bharadwaj, S., Gupta, A. K., & Sahu, A. K. (2024) "Characterisation of Contaminated Soil and Groundwater Due To Landfill Leachate" In Indian Geotechnical Conference (IGC 2024) Geotechnical Engineering for a Sustainable Tomorrow (GEST 2024), Aurangabad, Maharashtra in 19-21 Dec 2024.
- Bharadwaj, S., Gupta, A. K., & Sahu, A. K. (2025) "Landfill Mining and Waste Management: Challenges, Opportunities, And Innovations" ASCE International Conference, Challenges and Innovations for sustainable smart cities (CISSC 2025), from 7-9 February 2025, Organised by PEC, Chandigarh.

TABLE OF CONTENTS

CANDIDATE'S DECLARATION iii

PLAGIARISM VERIFICATION v

ABSTRACT vi

LIST OF PUBLICATIONS ix

CHAPTER 1: INTRODUCTION 1

1.1 GENERAL 1

1.2 PROBLEM STATEMENT 3

1.3 RESEARCH GAPS 5

1.4 MOTIVATION OF RESEARCH 7

1.5 RESEARCH OBJECTIVES 8

1.6 SCOPE OF RESEARCH AND RESEARCH METHODOLOGY OVERVIEW 9

1.7 ORGANIZATION OF THE THESIS 12

CHAPTER 2: LITERATURE REVIEW 14

2.1 INTRODUCTION 14

2.2 LANDFILL LEACHATE CHARACTERISTICS 16

2.3 IMPACT OF LEACHATE ON SOIL PROPERTIES 19

2.4 GROUNDWATER CONTAMINATION DUE TO LANDFILL LEACHATE 24

2.5 APPLICATION OF MACHINE LEARNING IN GROUNDWATER QUALITY PREDICTION AND SOIL HEAVY METAL PREDICTION 27

2.6 SYNTHESIS OF LITERATURE REVIEW AND RELEVANCE TO THE PRESENT STUDY 36

CHAPTER 3: PHYSICOCHEMICAL, HEAVY METAL AND GEOTECHNICAL ASSESSMENT OF LANDFILL LEACHATE-IMPACTED SOIL 37

3.1 INTRODUCTION 37

3.2 MATERIALS AND METHODS 39

3.2.1 STUDY AREA DESCRIPTION 39

3.2.2 SAMPLE COLLECTION AND PRESERVATION 41

3.2.3 LABORATORY ANALYSIS 44

3.2.4 DATA ANALYSIS AND POLLUTION ASSESSMENT METHODS 47

3.3 LANDFILL LEACHATE CHARACTERIZATION 49

3.3.1 PHYSICOCHEMICAL AND HEAVY METAL CHARACTERISTICS OF LEACHATE 50

3.3.2 LEACHATE POLLUTION INDEX ASSESSMENT 52

3.4 PHYSICOCHEMICAL AND HEAVY METAL ASSESSMENT OF LEACHATE-IMPACTED SOIL 55

3.4.1 PHYSICOCHEMICAL AND HEAVY METAL CHARACTERISTICS OF SOIL 55

3.4.2 SPATIAL AND VERTICAL DISTRIBUTION OF HEAVY METALS IN SOIL 61

3.4.3 STATISTICAL ANALYSIS OF HEAVY METAL CONTAMINATION 63

3.4.4 GEO-ACCUMULATION INDEX AND ENRICHMENT FACTOR ASSESSMENT 67

3.5 GEOTECHNICAL BEHAVIOUR OF LEACHATE-IMPACTED SOIL 73

3.5.1 INDEX AND COMPACTION PROPERTIES 74

3.5.2 PERMEABILITY AND SHEAR STRENGTH CHARACTERISTICS 83

3.6 SEM–EDX-BASED MICROSTRUCTURAL AND ELEMENTAL ANALYSIS OF SOIL 86

3.7 CONCLUSIONS 102

CHAPTER 4: HYDROGEOCHEMICAL ANALYSIS OF GROUNDWATER CONTAMINATION DUE TO LANDFILL LEACHATE 107
4.1 INTRODUCTION 107
4.2 MATERIALS AND METHODS 109
4.2.1 STUDY AREA AND SAMPLING LOCATIONS 109
4.2.2 GROUNDWATER AND SURFACE WATER SAMPLING 110
4.2.3 WEIGHTED ARITHMETIC WATER QUALITY INDEX (WA-WQI) 113
4.3 RESULTS AND DISCUSSION OF GROUNDWATER QUALITY 114
4.3.1 PHYSICOCHEMICAL ANALYSIS FOR BOREHOLE SAMPLES 114
4.3.2 HEAVY METAL CONTAMINATION FOR BOREHOLE SAMPLES 116
4.3.3 IDENTIFICATION OF CONTAMINATION IN HAND PUMP WATER SAMPLES 119
4.3.4 IDENTIFICATION OF CONTAMINATION IN SUBMERSIBLE PUMP WATER SAMPLES 121
4.3.5 IDENTIFICATION OF CONTAMINATION IN SURFACE WATER SAMPLE 123
4.4 WATER QUALITY INDEX (WQI) FOR ALL WATER SAMPLES 125
4.5 CONCLUSION 126
CHAPTER 5: DEVELOPMENT OF PREDICTIVE MODELLING FRAMEWORK FOR CONTAMINANT MIGRATION 128
5.1 INTRODUCTION 128
5.2 MATERIALS AND METHODS 130
5.2.1 STUDY AREA 130
5.2.2 GROUNDWATER SAMPLES 132
5.2.3 SOIL SAMPLING 135
5.2.4 WEIGHTED ARITHMETIC WATER QUALITY INDEX 136
5.2.5 MACHINE LEARNING MODELS 137
5.2.6 RESAMPLING TECHNIQUES AND HYPERPARAMETER OPTIMIZATION 141

5.2.7 DATA ANALYSIS AND FEATURE SELECTION 142
5.2.8 MODEL EVALUATION METRICS 143
5.2.9 TAYLOR DIAGRAM AND CONVERGENCE ANALYSIS 146
5.3 RESULTS AND DISCUSSION 147
5.3.1 ASSESSMENT OF GROUNDWATER QUALITY USING PHYSICOCHEMICAL PARAMETERS 147
5.3.2 GEOSPATIAL ANALYSIS OF PHYSICO-CHEMICAL ATTRIBUTES AND EVALUATED WQI 151
5.3.3 PERFORMANCE ANALYSIS OF MACHINE LEARNING MODELS TO PREDICT WQI 153
5.3.4 ANALYSIS OF MODEL PERFORMANCE UTILIZING ALL 7 FEATURES BASED ON PEARSON CORRELATION COEFFICIENT TO PREDICT WQI 156
5.3.5 ANALYSIS OF MODEL PERFORMANCE USING TOP 4 FEATURES BASED ON PEARSON CORRELATION COEFFICIENT TO PREDICT WQI 161
5.3.6 ANALYSIS OF MODEL PERFORMANCE USING TOP 2 FEATURES BASED ON PEARSON CORRELATION COEFFICIENT TO PREDICT WQI 164
5.3.7 ANALYSIS OF MODEL PERFORMANCE USING TOP MOST FEATURE BASED ON PEARSON CORRELATION COEFFICIENT TO PREDICT WQI 166
5.3.8 FURTHER PERFORMANCE CHECK AND VALIDATION ON THE BEST ML MODELS 168
5.3.9 ASSESSMENT OF SOIL PHYSICOCHEMICAL PROPERTIES AND HEAVY METAL CONTAMINATION 172
5.3.10 PERFORMANCE ANALYSIS OF MACHINE LEARNING MODELS FOR SOIL SAMPLES . 174
5.3.11 MODEL VALIDATION FOR SOIL SAMPLES 175
5.3.12 CONVERGENCE ANALYSIS OF THE OPTIMIZED MODEL 176
5.4 CONCLUSION 178
CHAPTER 6: CONCLUSION, FUTURE SCOPE AND SOCIAL IMPACT 180
6.1 GENERAL 180
6.2 CONCLUSION 180
6.3 SOCIAL IMPACT 182

6.4 FUTURE SCOPE 184

REFERENCES 185

PLAGIARISM REPORT 215

BRIEF CURRICULUM VITAE 216

LIST OF TABLES

Table 2.1 Summary of Literature on Landfill Leachate Characteristics and Pollution Potential.....18

Table 2.2 Summary of Literature on Leachate-Induced Changes in Soil Properties 22

Table 2.3 Summary of Literature on Groundwater Contamination due to Landfill Leachate.....26

Table 2.4 Summary of Comparative analysis of WQI for ML MODELS.....30

Table 2.5 Summary of key machine learning studies related to soil heavy metal prediction.....34

Table 3.1 Methods Used for Leachate Sample Analysis.....45

Table 3.2 Instruments/Methods Used for Physicochemical and Heavy Metal Analysis of Soil Samples 46

Table 3.3 Physicochemical and Heavy Metal Characteristics of Leachate Samples Collected from Bhalswa Landfill 51

Table 3.4 Leachate pollution Index.....54

Table 3.5 Statistical Analysis of Heavy Metals Concentration in soil samples 64

Table 3.6 Igeo Classification (Müller, 1969).....68

Table 3.7 Descriptive Summary for Geo-Accumulation Index (Igeo) 69

Table 3.8 Descriptive Summary for Enrichment Factor 71

Table 3.9 Enrichment Factor Classification 72

Table 3.10 Cross-Borehole EDX Summary.....86

Table 4.1 Physicochemical analysis of GW samples from borehole.....116

Table 4.2 Heavy metal analysis of GW samples from borehole 118

Table 4.3Physicochemical analysis of GW samples from handpumps 120

Table 4.4 Physicochemical analysis of GW samples from submersible pump samples 122

Table 4.5 Physicochemical analysis of surface water samples.....124

Table 4.6 Water quality indexing classification for all groundwater samples 125

Table 5.1 Class-wise Distribution of WQI.....148

Table 5.2 Statistical Summary of Geochemical Parameters from 488 Groundwater Samples after Water Quality Index Assessment.....148

Table 5.3 Model architecture obtained via hyperparameter tuning (All 7 Features).....155

Table 5.4 Performance metrics of ML models using all 7 features.....157

Table 5.5 Models Ranked on basis of distance from reference point in Taylor diagram.....160

Table 5.6 Performance Metrics of ML Models Using Top 4 Features (EC, Mg, K, TH) 161

Table 5.7 Models Ranked on Basis of Distance from Reference Point in Taylor Diagram (Top 4 Features) 163

Table 5.8 Performance Metrics of ML Models Using Top 2 Features (Mg, K).....164

Table 5.9 Models Ranked on Basis of Distance from Reference Point in Taylor Diagram (Top 2 Features) 166

Table 5.10 Performance Metrics of ML Models Using Single Feature (K).....166

Table 5.11 Models Ranked on Basis of Distance from Reference Point in Taylor Diagram (Single Feature K).....168

Table 5.12 Best Performing Models in Different Feature Cases 169

Table 5.13 Performance Metrics of Validated Models on Independent 70-Sample Dataset.....169

Table 5.14 Models Ranked on Basis of Distance from Reference Point in Taylor Diagram Final Validation.....171

Table 5.15 Statistical summary of soil physicochemical properties and heavy metals 173

Table 5.16 Soil model comparison with feature selection and optimization 176

Table 5.17 Cost function values at the 40th iteration.....178

Table 6.1 contribution of research work to sustainable development goals (SDGs).....183

LIST OF FIGURES

Figure 1.1 Bhalswa Dumpsite, Delhi, India 5

Figure 1.2 Flow chart of the research methodology adopted in the present study 11

Figure 3.1 Field view of MSW accumulation and visible leachate seepage at Bhalswa landfill, Delhi. 38

Figure 3.2 Location map showing soil sampling points in the vicinity of Bhalswa landfill, Delhi. 40

Figure 3.3 Leachate Collection Point at Bhalswa Landfill 41

Figure 3.4 Schematic representation of soil sampling profile 42

Figure 3.5 Schematic representation of boreholes distances from Bhalswa landfill, Delhi, India 43

Figure 3.6 Soil sampling points in the vicinity of Bhalswa landfill, Delhi, India 44

Figure 3.7 Laboratory testing (a) Saturation of Soil Samples (b) Orbital shaker with soil samples in flasks (c) Centrifuge machine with soil samples (d) Titration for total hardness (e) Performing test in lab (f) Volumetric titration for Chloride (g) Acid Digestion of Soil Sample (h) Prepared samples for testing 46

Figure 3.8 Leachate sample quality results..... 50

Figure 3.9 Depth-wise assessment of EC and pH in soil samples 56

Figure 3.10 Depth-wise assessment of Alkalinity, Cl, Ca and TH in soil samples 57

Figure 3.11 Depth-wise assessment of Mg and K in soil samples..... 57

Figure 3.12 Depth-wise assessment of Mn in soil samples..... 58

Figure 3.13 Depth-wise assessment of Cu, As, Ag and Cr in soil samples..... 59

Figure 3.14 Depth-wise assessment of Fe and Al in soil samples 59

Figure 3.15 Depth-wise assessment of Zn, Cd and Pb in soil samples..... 60

Figure 3.16 Distance-wise assessment of Alkalinity, Cl, Ca, TH assessment in soil samples..... 61

Figure 3.17 Distance-wise assessment of Mn, Mg, K in soil samples..... 62

Figure 3.18 Distance-wise assessment of Zn, Cd, Cu, As, Ag, Pb and Cr in soil samples..... 62

Figure 3.19 Distance-wise assessment of Fe and Al in soil samples..... 63

Figure 3.20 Spearman correlation heatmap showing relationships among heavy metals, sampling depth, and distance from landfill in 50 soil samples..... 66

Figure 3.21 (a) Particle Size distribution for BH-1..... 75

Figure 3.21 (b) Particle Size distribution for BH-2..... 75

Figure 3.21 (c) Particle Size distribution for BH-3..... 76

Figure 3.21 (d) Particle Size distribution for BH-4..... 76

Figure 3.21 (e) Particle Size distribution for BH-5..... 76

Figure 3.21 (f) Particle Size distribution for BH-6..... 77

Figure 3.21 (g) Particle Size distribution for BH-7..... 77

Figure 3.21 (h) Particle Size distribution for BH-8..... 77

Figure 3.21 (i) Particle Size distribution for BH-9..... 78

Figure 3.21 (j) Particle Size distribution for BH-10..... 78

Figure 3.22 Compaction Curves for BH-1 to BH-10..... 81

Figure 3.23 Direct Shear test plots for BH-1 to BH-10..... 84

Figure 3.24 EDS Spectrum — BH-1 (Highest Peak Area, EDAX APEX™)..... 88

Figure 3.25 SEM observation of BH-1 soil sample..... 88

Figure 3.26 EDS Spectrum — BH-2 (Highest Peak Area, EDAX APEX™)..... 89

Figure 3.27 SEM observation of BH-2 soil sample..... 89

Figure 3.28 EDS Spectrum — BH-3 (Highest Peak Area, EDAX APEX™)..... 91

Figure 3.29 SEM observation of BH-3 soil sample..... 91

Figure 3.30 EDS Spectrum — BH-4 (Highest Peak Area, EDAX APEX™)..... 92

Figure 3.31 SEM observation of BH-4 soil sample..... 92

Figure 3.32 EDS Spectrum — BH-5 (Highest Peak Area, EDAX APEX™)..... 94

Figure 3.33 SEM observation of BH-5 soil sample 94

Figure 3.34 EDS Spectrum — BH-6 (Highest Peak Area, EDAX APEX™)..... 95

Figure 3.35 SEM observation of BH-6 soil sample 95

Figure 3.36 EDS Spectrum — BH-7 (Highest Peak Area, EDAX APEX™)..... 97

Figure 3.37 SEM observation of BH-7 soil sample 97

Figure 3.38 EDS Spectrum — BH-8 (Highest Peak Area, EDAX APEX™)..... 98

Figure 3.39 SEM observation of BH-8 soil sample 98

Figure 3.40 EDS Spectrum — BH-9 (Highest Peak Area, EDAX APEX™)..... 100

Figure 3.41 SEM observation of BH-9 soil sample 100

Figure 3.42 EDS Spectrum — BH-10 (Highest Peak Area, EDAX APEX™)..... 102

Figure 3.43 SEM observation of BH-10 soil sample..... 102

Figure 4.1 Groundwater Sampling location in the vicinity of Bhalswa landfill, Delhi. 109

Figure 4.2 (a) GW sample collection from Handpump (b) Bhalswa lake (c) GW samples from different sources (d) GW sample collection from Borehole location in the vicinity of Bhalswa landfill, Delhi. 111

Figure 4.3 Physicochemical characteristics of borehole (BH1-BH9) groundwater samples of (a) EC, TDS, Major Ions, Hardness, and (b) pH. 115

Figure 4.4 Heavy metal and major element concentrations in groundwater samples. 117

Figure 4.5 Physicochemical characteristics of hand pump (HP1-HP24) groundwater samples of (a) TDS, Major Ions, Hardness, and (b) EC, pH. 119

Figure 4.6 Physicochemical characteristics of submersible (SM1-SM13) groundwater samples of (a) EC, TDS, Major Ions, Hardness, and (b)pH. 122

Figure 4.7 Physicochemical Characteristics of Surface Water Samples from Bhalswa Lake.. 124

Figure 4.8 Weighted Arithmetic-Water Quality Index (WA-WQI) for borehole, hand pump, and submersible. 125

Figure 5.1 Ground water sample locations (Delhi, NCR (National Capital Region), India).... 131

Figure 5.2 Study location Map (a) India (b) Delhi (c) North District Delhi (d) Bore Hole locations for soil sampling 132

Figure 5.3 Flowchart for machine learning-based prediction of water quality..... 134

Figure 5.4 Flowchart of steps for developing WA-WQI 136

Figure 5.5 Distribution of geochemical water quality parameters used for machine learning models from 488 samples in Delhi and its surrounding region, India..... 149

Figure 5.6 Pearson Correlation Matrix Heatmap for all 7 features and WQI 150

Figure 5.7 Spatial Distribution map showing geospatial variation of geochemical attributes (pH, EC, Cl, Mg, Ca, K, TH) and WQI in Delhi NCR, India. 152

Figure 5.8 Comparison of R^2 scores and MAPE for various models in training over k-fold values..... 155

Figure 5.9 Bar graphs showing comparison among ML models on basis of different metrics 159

Figure 5.10 Scatter plot Actual Vs Predicted WQI for various ML models (a) ANN (b) SVM, (C) XGBoost (d) RF, (e) KNN (f) DT..... 158

Figure 5.11 Taylor diagram using all 7 features 160

Figure 5.12 Line plot of Actual vs. Predicted WQI for (a) KNN, (b) SVM, (c)DT, (d)RF, (e)XGBOOST, (f)ANN using top 4 features..... 162

Figure 5.13 Taylor diagram (using top 4 features)..... 163

Figure 5.14 Line plot of Actual vs. Predicted WQI for (a) KNN, (b)SVM, (c)DT, (d)RF, (e)XGBOOST, (f)ANN using top 2 features..... 165

Figure 5.15 Taylor diagram (using top 2 features.)..... 165

Figure 5.16 Line plot of Actual vs. Predicted WQI for (a)KNN, (b)SVM, (C)DT, (d)RF, (e)XGBOOST, (f)ANN using top most features (K)..... 167

Figure 5.17 Taylor diagram (using top most feature 'K') 168

Figure 5.18 Scatter plot of Predicted vs Actual WQI of ANN (Case-1)..... 170

Figure 5.19 Scatter plot of Predicted vs Actual WQI of SVM (Case-2) 170

Figure 5.20 Scatter plot of Predicted vs Actual WQI of SVM (Case-3) 171

Figure 5.21 Scatter plot of Predicted vs Actual WQI of SVM (Case-4) 171

Figure 5.22 Taylor diagram for further performance check..... 172

Figure 5.23 (a)-(g) Illustrates various model evaluation analysis using developed approach over basic models, (a) Accuracy (%), (b) Precision (%), (c) Recall (%), (d) Sensitivity (%), (e) Specificity (%), (f) False Negative Rate (%)..... 175

Figure 5.24 Illustrates convergence analysis of developed approach over compared approaches 177

LIST OF ABBREVIATION

MSW	Municipal Solid Waste
UNSDG	United Nations Sustainable Development Group
HDPE	High-Density Polyethylene
MP-AES	Microwave Plasma Atomic Emission Spectroscopy
I_{geo}	Geo-Accumulation Index
EF	Enrichment Factor
LPI	Leachate Pollution Index
WHO	World Health Organization
FAO	Food And Agriculture Organization
EDTA	Ethylenediaminetetraacetic Acid
DO	Dissolved Oxygen
PH	Potential Of Hydrogen
Cl	Chloride
EC	Electrical Conductivity
Ca	Calcium
TH	Total Hardness
TSS	Total Suspended Solid
TDS	Total Dissolved Solids
COD	Chemical Oxygen Demand
CV	Coefficient Of Variation
P	Spearman Correlation Coefficient
SP	Poorly Graded Sand
SP-SM	Poorly Graded Sand With Silt
SM	Silty Sand With Gravel
SC	Clayey Sand With Gravel
CL	Low-Plasticity Clay
PSD	Particle Size Distribution
OMC	Optimum Moisture Content
MDD	Maximum Dry Density
SEM	Scanning Electron Microscopy
EDX	Energy-Dispersive X-Ray Spectroscopy
WA-WQI	Weighted Arithmetic Water Quality Index
CGWB	Central Ground Water Board
CPCB	Central Pollution Control Board
BH	Bore-Hole
SM	Submersible
HP	Handpump
GW	Groundwater
ML	Machine Learning

ANN	Artificial Neural Network
RF	Random Forest
KNN	K-Nearest Neighbors
DT	Decision Tree
XGBOOST	Extreme Gradient Boosting
SVM	Support Vector Machine
RF-GA	Random Forest And Genetic Algorithm
LR	Logistic Regression
R^2	Coefficient Of Determination
MSE	Mean Squared Error
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
RMSE	Root Mean Square Error
CRMSE	Centered Root Mean Square Error
IDW	Inverse Distance Weighting
ARGs	Antibiotic Resistance Genes
ARB	Antibiotic-Resistant Bacteria
MRGs	Metal Resistance Genes

CHAPTER 1: INTRODUCTION

1.1 GENERAL

Rapid urbanization, population growth and changing consumption patterns have increased municipal solid waste generation across the world. In many developing countries, waste management systems have not expanded at the same pace as waste generation, and open dumping or poorly engineered landfill disposal continues to remain common. Such disposal sites often lack adequate liners, leachate collection systems, gas control arrangements and scientific containment measures, which increases the risk of contaminant movement into the surrounding environment (Ahmed et al., 2020; Kumar & Alappat, 2005; Hussein et al., 2021). Landfill leachate is one of the major pathways through which waste disposal sites affect environmental quality. It is produced when rainfall, moisture and waste decomposition processes interact with dumped municipal solid waste, generating a contaminated liquid containing organic matter, inorganic salts, nitrogenous compounds, chlorides, heavy metals and microbial pollutants (Kumar & Alappat, 2005; Rani et al., 2020). The pollution potential of leachate is important because it can migrate into soil, surface water and groundwater depending on site conditions, rainfall, soil permeability, landfill age and local hydrogeology (Ahamad et al., 2019; Kenue et al., 2023; Saha et al., 2023).

In India, the problem is more serious because many waste disposal sites have historically functioned as open or non-engineered dumps. Studies on Indian landfill and dumpsite environments have reported leachate-related contamination in soil and groundwater, including heavy metal accumulation, high salinity, poor groundwater quality and potential human health risks (Kanmani & Gandhimathi, 2013; Ahamad et al., 2019; Drall et al., 2025). Delhi represents a critical urban case because its major landfill sites have received large quantities of municipal solid waste over long periods, and previous studies have reported concerns related to leachate quality, groundwater deterioration, landfill fires and surrounding environmental risk (Ahmed et al., 2020; Sharma et al., 2024)

The soil and groundwater are the two main environmental media affected by landfill leachate. The soil acts as the first receiving medium and may undergo changes in pH, electrical conductivity, salinity, nutrient content, heavy metal concentration, microbial activity and enzyme activity due to leachate interaction (Kanmani & Gandhimathi, 2013; Islam et al., 2024; Kooch et al., 2023). Leachate may also alter the engineering behaviour of soil by affecting Atterberg's limits, compaction characteristics, permeability and shear strength, which is important for understanding contaminant migration and ground stability near landfill sites (Harun et al., 2013; Khodary et al., 2018; Bharadwaj et al., 2025a). Once contaminants move beyond the soil's retention capacity, they may enter groundwater and affect its suitability for domestic or other uses (Srivastava & Ramanathan, 2008; Ahamad et al., 2019).

The Bhalswa landfill in Delhi is a suitable site for examining these processes because it is an old municipal solid waste disposal site located in a densely inhabited urban setting. Previous Bhalswa-focused studies have reported that the landfill lacks scientific lining and that leachate, soil and groundwater samples from the surrounding area show changes in physicochemical characteristics (Bharadwaj et al., 2025a; Bharadwaj et al., 2025b). Groundwater studies around Bhalswa have also reported elevated electrical conductivity, major ions, heavy metals and health-risk concerns, indicating the need for continued assessment of the landfill-affected zone (Srivastava & Ramanathan, 2008; Ahamad et al., 2019; Central Ground Water Board [CGWB], 2024).

The present research is developed in this context and focuses on changes in soil and groundwater behaviour due to contamination from landfill leachate in the vicinity of Bhalswa landfill. The study considers leachate as the contamination source, soil as the receiving and controlling medium, and groundwater as the affected resource. In addition to field and laboratory-based assessment, the research incorporates machine learning for groundwater quality prediction and soil heavy metal prediction, as data-driven models can help interpret complex environmental datasets with multiple

interacting parameters (Bharadwaj et al., 2025c; Bharadwaj et al., 2025d; Jia et al., 2019; Yang et al., 2021).

The study is also aligned with the United Nations Sustainable Development Goals (UNSDG). Protection of groundwater quality supports SDG 6 on clean water and sanitation, while reducing exposure to contaminated soil and groundwater supports SDG 3 on good health and well-being. Scientific landfill management and reduction of open dumping support SDG 11 on sustainable cities and communities and SDG 12 on responsible consumption and production. Soil protection and prevention of contaminant movement into ecosystems further support SDG 15 on life on land (United Nations, 2015).

This chapter is organised as follows: **Section 1.2** presents the problem statement of the study; **Section 1.3** discusses the research gaps; **Section 1.4** presents the motivation of the research; **Section 1.5** outlines the research objectives; **Section 1.6** describes the scope of the research and methodology overview; and **Section 1.7** presents the organization of the thesis.

1.2 PROBLEM STATEMENT

Municipal solid waste disposal through open dumping and poorly engineered landfill practices has become a persistent environmental problem in many Indian cities. In such systems, waste is often deposited without adequate segregation, bottom lining, leachate collection, gas control or long-term containment measures, allowing leachate to migrate into surrounding soil and groundwater. This creates a continuing contamination pathway in which the landfill behaves not only as a waste storage site, but also as a long-term source of subsurface pollution (Kumar & Alappat, 2005; Ahmed et al., 2020; Drall et al., 2025).

The severity of this issue increases in large urban centres where landfill sites are located close to densely populated settlements. Delhi represents a critical case because large quantities of

municipal solid waste are disposed of at major landfill sites such as Bhalswa, Okhla and Ghazipur. Among these, Bhalswa landfill (Figure 1.1.), is of particular concern because it was commissioned in 1994, receives about 2500 MT/day of solid waste, covers about 52 acres and has developed into a waste mound of about 60–65 m above the surrounding ground level (CGWB, 2024). The site has also been described as non-engineered and unlined, which increases the possibility of leachate seepage into the surrounding subsurface environment (Ahamad et al., 2019; Bharadwaj et al., 2025a).

The environmental problem is not limited to groundwater quality alone. As leachate passes through the soil profile, it can alter soil physicochemical properties such as pH, electrical conductivity, salinity, alkalinity, nutrient content and heavy metal concentration (Kanmani & Gandhimathi, 2013; Islam et al., 2024). Leachate contamination may also affect the engineering behaviour of soil, including permeability, compaction characteristics, moisture response and shear strength, which can influence contaminant transport and ground stability near landfill sites (Harun et al., 2013; Khodary et al., 2018; Bharadwaj et al., 2025b).

Previous field-based studies around Bhalswa have reported elevated electrical conductivity, chloride, nitrate, fluoride, heavy metals and other hydrochemical indicators in groundwater, suggesting the influence of leachate migration from the landfill (Srivastava & Ramanathan, 2008; Ahamad et al., 2019; CGWB, 2024). Site-specific investigations have also shown that several soil and groundwater parameters near Bhalswa exceed permissible or desirable limits, while soil and groundwater quality tend to improve with increasing distance from the landfill (Bharadwaj et al., 2025a). These findings indicate that Bhalswa landfill affects the surrounding environment through a connected soil–groundwater pathway rather than through isolated contamination of a single medium.

Thus, the core problem addressed in the present study is the lack of an integrated understanding of how landfill leachate changes both soil and groundwater behaviour in the vicinity of Bhalswa

landfill. The landfill acts as the contaminant source, soil acts as the first receiving and controlling medium, and groundwater becomes the affected resource after subsurface migration. In the absence of effective containment, treatment and regular monitoring, this process can continue over time and may pose long-term environmental and health risks to nearby populations (Sidhu et al., 2015; Ahamad et al., 2019; CGWB, 2024).



Figure 1.1 Bhalswa Dumpsite, Delhi, India

1.3 RESEARCH GAPS

Earlier studies have established that landfill leachate can degrade groundwater quality, contaminate soil with heavy metals, and create ecological and human health risks near open or unengineered landfill sites. Studies around Bhalswa and other Indian dumpsites have reported groundwater quality deterioration, heavy metal contamination, and the need for better monitoring and remediation (Srivastava & Ramanathan, 2008; Ahamad et al., 2019; Drall et al., 2025). However, most studies remain focused on groundwater quality, while soil behaviour is

comparatively less explored, especially in terms of physicochemical changes, geotechnical response, and contaminant movement through the subsurface.

A major gap is the limited investigation of how contamination varies both horizontally with distance from the landfill and vertically with depth. Since soil acts as the first receiving and transmitting medium for leachate, understanding depth-wise and distance-wise variation is important for interpreting pollutant migration. Existing studies also provide limited integration of leachate characteristics, soil contamination, groundwater quality, surface water impact, and predictive modelling within a single framework. This limits the understanding of landfill contamination as a connected source–pathway–receptor system.

Another important gap is the limited use of validated machine learning tools for landfill-related contamination assessment. Although machine learning has been used for groundwater quality prediction, fewer studies connect predictive modelling with field-based soil and groundwater contamination around landfill sites. Prediction of soil heavy metals and groundwater quality can help reduce repeated laboratory burden and support monitoring, but such tools require proper validation and site-specific interpretation (Bharadwaj et al., 2025c; Bharadwaj et al., 2025d).

Based on these gaps, the present study focuses on an integrated assessment of landfill-leachate-induced changes in soil and groundwater behaviour around Bhalswa landfill. The study addresses soil contamination, horizontal and vertical variation, geotechnical behaviour, groundwater quality, and machine learning-based prediction within a single research framework. This directly supports a more complete understanding of contaminant movement and environmental risk in the vicinity of Bhalswa landfill.

Based on the literature review as mentioned in **Chapter 2**, the key gaps addressed in the present study are the following:

a. Lack of detailed soil-focused assessment in landfill leachate studies, especially in the Bhalswa region.

b. Lack of systematic horizontal and vertical investigation of contaminated soil characteristics.
c. Limited attention to the geotechnical behaviour of soil affected by landfill leachate.
d. Absence of an integrated framework combining soil, leachate, groundwater, surface water, and pollution assessment.
e. Insufficient development and validation of predictive modelling tools for landfill-related contaminant assessment.
f. Need for stronger decision-support tools that reduce laboratory burden while preserving monitoring reliability.

1.4 MOTIVATION OF RESEARCH

The motivation for the present research arises from the need to understand landfill contamination as a connected soil–groundwater system. In landfill-affected areas, leachate originates from the waste mass, interacts with soil as the first receiving medium, and may subsequently migrate towards groundwater depending on soil permeability, stratification, rainfall infiltration, contaminant strength and local hydrogeological conditions (Kumar & Alappat, 2005; Kenue et al., 2023). Therefore, an integrated assessment is required to examine how leachate modifies soil properties and how these changes influence groundwater contamination.

Previous studies have provided important evidence on leachate pollution, groundwater deterioration, heavy metal contamination and health risks near landfill sites (Srivastava & Ramanathan, 2008; Ahamad et al., 2019; Alam et al., 2020; Drall et al., 2025). However, for a complete understanding of landfill-affected environments, contamination levels need to be studied together with spatial variation, soil behaviour and contaminant migration pathways. This is important because soil does not only receive contaminants, but also controls their retention,

transformation and movement towards groundwater (Kanmani & Gandhimathi, 2013; Khodary et al., 2018; Islam et al., 2024).

Bhalswa landfill provides a significant case for such an assessment because it has a long operational history, high waste load, absence of scientific lining and close proximity to residential areas. Earlier studies around Bhalswa have reported leachate impact, groundwater quality deterioration and potential health risks, while recent site-specific work has also examined soil, groundwater and leachate characteristics around the landfill (Ahamad et al., 2019; Bharadwaj et al., 2025a; Bharadwaj et al., 2025b; CGWB, 2024). These findings motivate a more detailed investigation of horizontal and vertical variation in soil properties and their connection with groundwater quality.

Another motivation for the study is the need for efficient and reliable monitoring tools. Conventional laboratory testing is essential, but repeated testing across multiple locations and parameters can be time-consuming and resource-intensive. Machine learning offers a useful supporting approach for predicting groundwater quality and soil heavy metal contamination using measured environmental parameters, thereby helping to reduce monitoring burden while improving interpretation of complex datasets (Bharadwaj et al., 2025c; Bharadwaj et al., 2025d; Jia et al., 2019; Yang et al., 2021). Therefore, the present research integrates field investigation, laboratory analysis and predictive modelling to support better understanding and management of landfill-leachate-induced soil and groundwater contamination around Bhalswa landfill.

1.5 RESEARCH OBJECTIVES

The entire work in the present thesis is divided into six chapters based on the following objectives listed below:

a. To investigate the change in physicochemical characteristics of soil due to contamination from landfill leachate in horizontal and vertical directions.
b. To investigate the change in soil behavior and properties due to contamination.
c. To develop and validate a predictive modelling framework for contaminant migration.
d. To investigate the extent of contamination of groundwater due to landfill leachate.

1.6 SCOPE OF RESEARCH AND RESEARCH METHODOLOGY OVERVIEW

The scope of the present research is limited to the assessment of landfill-leachate-induced changes in soil and groundwater behaviour in the vicinity of Bhalswa landfill, Delhi. The study considers landfill leachate as the contaminant source, soil as the first receiving and controlling medium, and groundwater as the affected resource. The work focuses on physicochemical contamination, heavy metal distribution, geotechnical response of soil, groundwater quality assessment, and machine learning-based prediction of groundwater quality and soil heavy metals. Bhalswa landfill is selected because it is a long-operating, non-engineered landfill located close to residential areas, and previous studies have reported leachate impact, groundwater deterioration and soil contamination concerns around the site (Srivastava & Ramanathan, 2008; Ahamad et al., 2019; Bharadwaj et al., 2025a; CGWB, 2024).

The research methodology follows an integrated field, laboratory and modelling-based approach. In the first stage, the study area is identified and sampling locations are selected around Bhalswa landfill to capture spatial variation with respect to distance from the landfill. Soil, groundwater and leachate samples are collected and analysed for relevant physicochemical parameters such as pH, electrical conductivity, total dissolved solids, chloride, alkalinity, hardness, calcium, magnesium, sulphate, nitrate and other selected parameters. Soil samples are further assessed for heavy metal contamination and geotechnical behaviour, because leachate interaction can influence

both chemical quality and engineering properties of soil (Kanmani & Gandhimathi, 2013; Harun et al., 2013; Khodary et al., 2018; Bharadwaj et al., 2025b).

Groundwater quality is assessed using measured hydrochemical parameters and standard water quality evaluation methods. Water Quality Index-based assessment is used to classify groundwater suitability and to understand the combined influence of multiple parameters on groundwater quality. The study also applies statistical and spatial interpretation to identify variation in contamination intensity across the study area. Such an approach is important because landfill-affected groundwater is usually influenced by leachate strength, soil permeability, depth to groundwater, local flow conditions and seasonal infiltration (Ahamad et al., 2019; Sharma et al., 2020; CGWB, 2024).

In the modelling stage, machine learning techniques are used to predict groundwater quality and soil heavy metal concentration from measured environmental parameters. Models such as Artificial Neural Network, Random Forest, Support Vector Machine, XGBoost, Decision Tree and related algorithms are compared using performance indicators such as R^2, RMSE, MAE and MAPE. This modelling component is included to support faster and cost-effective environmental monitoring, since repeated laboratory testing of large numbers of samples can be time-consuming and expensive. Previous work on Delhi-NCR groundwater datasets has shown that machine learning can effectively predict Water Quality Index values and support groundwater quality monitoring (Bharadwaj et al., 2025c; Bharadwaj et al., 2025d).

Thus, the methodology combines contaminant source characterization, soil assessment, groundwater quality evaluation and predictive modelling within one framework. The overall methodological framework adopted in the present study is presented in Figure 1.2.

The research does not treat soil and groundwater as isolated media; instead, it examines how leachate moves through the subsurface and contributes to changes in both soil and groundwater

behaviour. This integrated scope provides the basis for understanding contaminant migration, identifying vulnerable zones and supporting improved environmental monitoring around Bhalswa landfill.

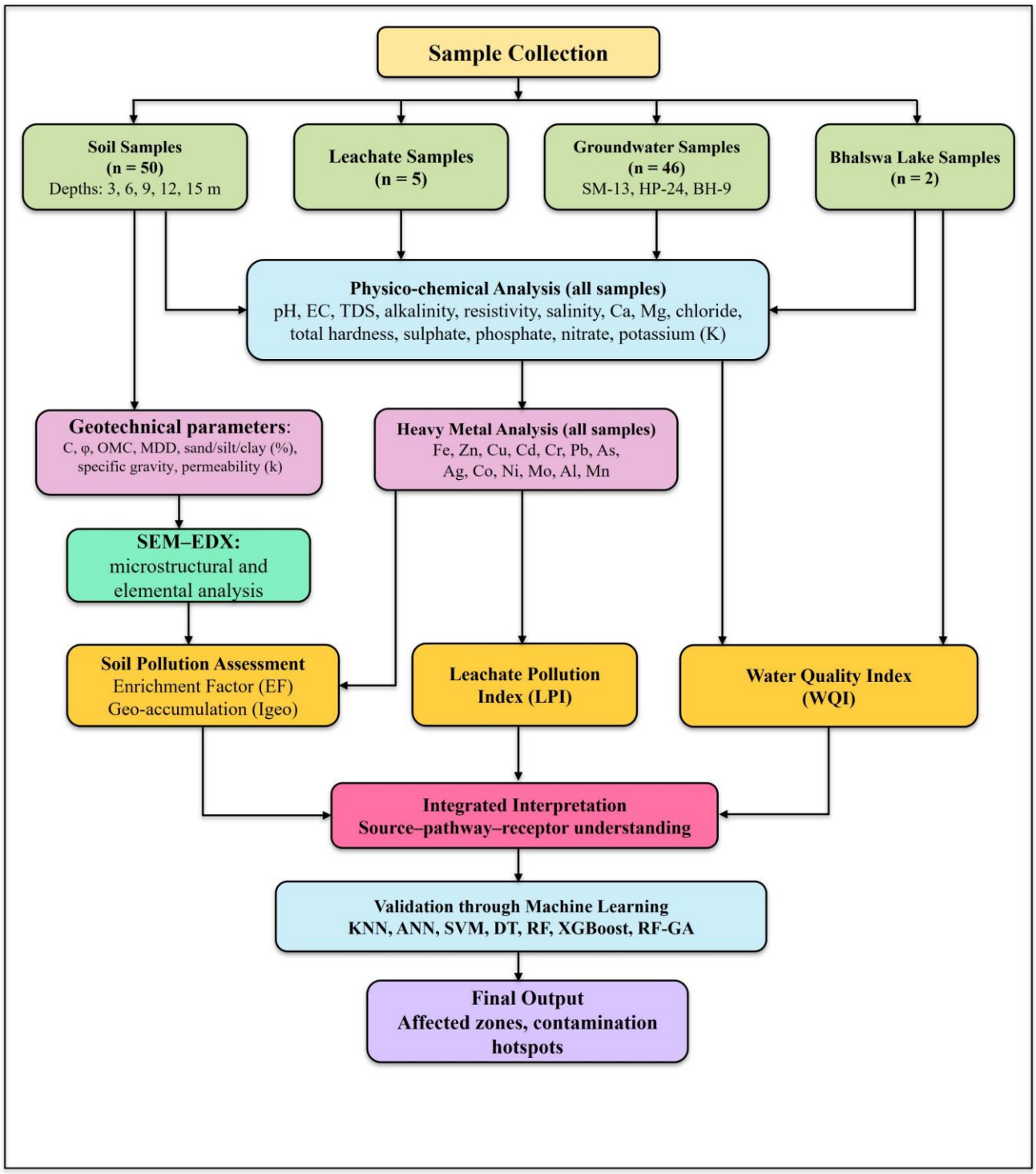


Figure 1.2 Flow chart of the research methodology adopted in the present study

1.7 ORGANIZATION OF THE THESIS

The thesis is organized into six chapters to examine changes in soil and groundwater behaviour due to contamination from landfill leachate in the vicinity of Bhalswa landfill.

Chapter 1: Introduction presents the background of the study, problem statement, research gaps, motivation, research objectives, scope of the research, and methodology overview. It establishes the need for an integrated assessment of landfill leachate, contaminated soil, groundwater quality, and predictive modelling.

Chapter 2: Literature review addresses existing studies related to landfill leachate characteristics, impact of leachate on soil properties, groundwater contamination due to landfill leachate, and the application of machine learning in groundwater quality and soil heavy metal prediction. This chapter provides the scientific basis for the present study.

Chapter 3: Physicochemical and Heavy Metal Characterization of Landfill Leachate and Contaminated Soil presents the analysis of landfill leachate and soil samples collected from the study area. It discusses physicochemical parameters, heavy metal concentrations, and changes in soil characteristics due to leachate contamination.

Chapter 4: Hydrogeochemical Analysis of Groundwater Contamination due to Landfill Leachate examines groundwater quality in the vicinity of Bhalswa landfill. It includes hydrogeochemical characterization, water quality assessment, heavy metal analysis, and interpretation of groundwater contamination caused by leachate migration.

Chapter 5: Development of Predictive Modelling Framework for Groundwater Quality and Soil Heavy Metal Contamination presents the machine learning-based prediction framework developed in the study. It includes groundwater quality prediction, soil heavy metal prediction, model training, validation, performance comparison, and interpretation of model outputs.

Chapter 6: Conclusions, Future Scope and Social Impact summarizes the major findings of the research and presents the conclusions drawn from the study. It also discusses the future scope of work and the social relevance of the research in terms of landfill monitoring, groundwater protection, soil contamination control, and public health risk reduction.

CHAPTER 2: LITERATURE REVIEW

2.1 INTRODUCTION

Municipal solid waste disposal remains a major environmental concern, especially in developing countries where open dumping and inadequately engineered landfills are still common. These disposal systems create long-term risks because contamination does not remain confined within the waste body, but gradually interacts with surrounding environmental media through subsurface pathways. Landfill-related pollution therefore requires an integrated review of waste disposal practices, soil response, groundwater vulnerability, and assessment methods rather than a narrow discussion of isolated contamination indicators (Kumar & Alappat, 2005; Sharma et al., 2020).

In the Indian context, the problem becomes more critical because many landfill sites have historically functioned as open or inadequately engineered dumps without effective liners, leachate collection systems and containment infrastructure. These conditions increase the possibility of leachate seepage into soil and its subsequent movement towards shallow groundwater systems (Rani et al., 2020; Central Ground Water Board [CGWB], 2024). Bhalswa landfill in Delhi represents this broader concern because it is one of the major landfill sites in the city, has been operational since 1994, and is located in an alluvial setting where subsurface material and groundwater conditions can influence contaminant movement (CGWB, 2024). The CGWB report records Bhalswa as a water-contaminated priority area and identifies groundwater quality, salinity, heavy metals and bacteriological contamination as concerns around the site (CGWB, 2024).

Existing landfill studies provide a wide base for understanding landfill-affected environments from different perspectives. Studies on leachate characteristics and pollution indices help assess the strength of the contaminant source, while groundwater-focused studies evaluate changes in water quality through physicochemical parameters, water quality indices, heavy metal indices and

spatial contamination patterns (Kumar & Alappat, 2005; Alam et al., 2020; Rani et al., 2020; Drall et al., 2025). Soil-focused studies add another important dimension by examining heavy metal accumulation, soil quality degradation, soil health indicators and geotechnical changes caused by leachate interaction (Islam et al., 2024; Kooch et al., 2024; Li et al., 2024; Mir et al., 2025). Recent studies have also expanded landfill pollution assessment by identifying emerging contaminants and microbial signatures in landfill leachate, showing the growing need for broader environmental evaluation (Qian et al., 2024; Robey et al., 2024; Yu et al., 2024). Building on these contributions, the present study examines Bhalswa landfill through a connected source, pathway, receptor perspective, where leachate is considered as the contaminant source, soil as the receiving and controlling medium, and groundwater as the affected resource. This approach is relevant because Bhalswa has strong leachate characteristics, brackish groundwater in the surrounding area and a subsurface profile consisting of sand, gravel, silt, clay and kankar, all of which can influence contaminant retention, migration and groundwater vulnerability (CGWB, 2024).

This Chapter is organized into the following sections: **Section 2.2** discusses landfill leachate characteristics; **Section 2.3** reviews the impact of leachate on soil properties; **Section 2.4** examines groundwater contamination due to landfill leachate; **Section 2.5** addresses the application of machine learning in groundwater quality prediction and soil heavy metal prediction; and **Section 2.6** presents a synthesis of literature review and relevance to the present study. Together, these sections are intended to move from source to pathway to impact, and finally to assessment and prediction. This structure reflects the overall logic of the thesis, in which landfill leachate is treated as the contamination source, soil and groundwater as the principal affected media, and predictive modelling as a tool for improved environmental evaluation and management.

2.2 LANDFILL LEACHATE CHARACTERISTICS

Landfill leachate is the liquid fraction generated when moisture percolates through municipal solid waste and extracts dissolved and suspended contaminants from decomposing waste material. The quality of leachate is not constant because landfill age, waste composition, rainfall, moisture content, hydrology, microbial degradation and operational conditions influence its chemical strength and pollutant composition (Kumar & Alappat, 2005; Rani et al., 2020). Therefore, leachate is generally assessed through a group of physicochemical, organic, inorganic, bacteriological and heavy-metal parameters rather than through a single contaminant indicator (Kumar & Alappat, 2005). In poorly engineered or unlined dump sites, this process becomes even more critical because the leachate is allowed to migrate directly into the subsurface without proper interception, collection, or treatment. For this reason, the characterization of landfill leachate forms a basic starting point in understanding landfill-induced environmental contamination (Bharadwaj et al., 2025c).

Landfill leachate has been characterized by high concentrations of organic constituents, dissolved salts, acids, metals and microorganisms, while its composition may vary with biodegradable matter and time (Kumar & Alappat, 2005). The Leachate Pollution Index was proposed to represent this composite pollution potential through selected parameters such as BOD_5, COD, pH, TKN, ammoniacal nitrogen, total dissolved solids, chloride, cyanide, phenolic compounds, total coliform bacteria and heavy metals (Kumar & Alappat, 2005; Rani et al., 2020).

In the Delhi landfill study, Bhalswa leachate showed BOD_5 of 2948.7 mg/L and COD of 5216.5 mg/L, indicating a high organic load and considerable biodegradable organic matter (Rani et al., 2020). Recent global analysis also shows that leachate commonly contains both biodegradable and recalcitrant toxic compounds, and landfill type and waste composition strongly control leachate quality (Qian et al., 2024; Lott et al., 2025). The inorganic fraction of leachate is mainly represented by total dissolved solids, electrical conductivity, chloride, sulphate, alkalinity,

calcium, magnesium, sodium, potassium, nitrate and ammoniacal nitrogen. These parameters are important because they indicate salinity, ionic strength and the potential for contaminant migration into soil and groundwater (Rani et al., 2020; Saghi et al., 2024). The CGWB report also recorded very high electrical conductivity in Bhalswa leachate, with values around 39000 and 38845 µS/cm, along with exceedance of fluoride and nitrate in leachate samples (CGWB, 2024).

Heavy metals form another critical component of landfill leachate because these elements are persistent and may remain in the environment for long periods. In the Delhi landfill study, Bhalswa leachate contained Fe, Cu, Ni, Zn, Pb, Cr, Hg and As, with arsenic reported at 1.53 mg/L and iron at 10.78 mg/L (Rani et al., 2020). The CGWB report also recorded Fe, As, Cr and Ni above permissible limits in both Bhalswa leachate samples, indicating that metal contamination remains relevant for the site-specific assessment of Bhalswa landfill (CGWB, 2024). Recent studies have expanded the understanding of landfill leachate beyond conventional BOD, COD, TDS and metals. PFAS, pharmaceuticals, microplastics, phthalates and other emerging contaminants have been reported in landfill leachate, indicating that leachate is a chemically complex effluent with both conventional and emerging pollutant groups (Robey et al., 2024; Yu et al., 2024; GodvinSharmila et al., 2023; Song et al., 2023).

For the present study, the characteristics of Bhalswa landfill leachate are important because the reported values show high organic load, high dissolved solids, high chloride, high ammoniacal nitrogen, bacteriological contamination and selected metal exceedance (Rani et al., 2020; CGWB, 2024). Such leachate composition indicates strong pollution potential and provides the chemical basis for examining subsequent changes in soil and groundwater behaviour around the landfill. Therefore, Section 2.2 establishes the pollutant strength and composition of landfill leachate and further Table 2.1 summarizes the various landfill leachate characteristics and pollution potential

Table 2.1 Summary of Literature on Landfill Leachate Characteristics and Pollution Potential

S.No.	Author / Year	Study area / focus	Parameters studied	Major findings
1	Kumar and Alappat (2005)	Municipal landfill leachate pollution assessment using Leachate Pollution Index	BOD, COD, pH, TKN, NH_4-N, TDS, chlorides, phenols, coliforms and heavy metals	The study explained that landfill leachate contains organic constituents, metals, acids, dissolved salts and microorganisms. It also stated that leachate characteristics vary with composition, volume, biodegradable matter and time.
2	Rani et al. (2020)	Bhalswa, Okhla and Ghazipur landfill leachate, Delhi	pH, BOD_5, COD, TKN, NH_4-N, TDS, chlorides, phenols, cyanide, coliforms and metals	The study described landfill leachate as a mixture of ammonia, organic compounds, inorganic compounds, ligands, biological organisms and toxic metals. It also noted that leachate composition changes with hydrology, waste type, waste composition, climate and landfill age.
3	Sharma et al. (2020)	Non engineered dumpsites in Solan, Mandi, Sundernagar and Baddi, Himachal Pradesh	pH, TDS, BOD, COD, TKN, NH_4-N, chlorides, calcium, sulphate, Fe, Cu, Ni, Zn, Pb, Hg, As, CN and Cr	The study found that leachate contains organic, inorganic and heavy metal contaminants. It also noted that leachate characteristics are affected by waste type, hydrology, rainfall volume and landfill age.
4	Qian et al. (2024)	Global landfill leachate characteristics from 318 municipal landfill sites	COD, NH_4^+, SO_4^{2-}, Cl^-, K^+, Na^+, Ca^{2+} and microbiome composition	The study reported that COD and ammonium are among the most dominant leachate components globally, followed by salts such as sulphate, chloride, potassium, sodium and calcium. It also linked microbial communities with nitrogen and salt concentrations.
5	Saghi et al. (2024)	Municipal solid waste landfill leachates in Iranian metropolises	Physicochemical characteristics and pollution indices	The study highlighted persistent organic and hazardous contaminants in municipal landfill leachates and emphasized the need for effective leachate management.
6	Robey et al. (2024)	Municipal solid waste landfill leachate from Puerto Rico	PFAS, physicochemical parameters, leachate, treated leachate and gas condensate	The study analyzed 20 landfill leachate samples and quantified 92 PFAS compounds. A total of 51 PFAS were detected above quantitation limits, and mean total PFAS concentration was reported as 38,000 ng/L.
7	Yu et al. (2024)	Pharmaceutical pollution in municipal landfill leachates	Pharmaceuticals, regional differences and influencing factors	The study identified municipal solid waste landfills as significant sources of pharmaceuticals in the environment. It also reported that non antibiotic pharmaceuticals may pose higher risk than antibiotics in landfill leachates.
8	Song et al. (2023)	Molecular characterization of organic compounds during full scale landfill leachate treatment	Organic compounds, emerging contaminants, GC×GC-QTOF-MS, A^2/O-MBR-NF treatment	The study detected 17 emerging contaminants with potential hazards, including diisobutyl phthalate and 4-tert-butylphenol. It showed that landfill leachate contains chemically complex organic compounds at the molecular level.
9	GodvinSharmila et al. (2023)	Microplastics as emerging contaminants in landfill leachate	Microplastics, source, impact and remediation technologies	The study explained that plastics buried in landfills gradually break down into microplastics due to physical, chemical and biological processes, making landfill leachate a potential source of microplastic contamination.
10	Lott et al. (2025)	Comparison of leachates by landfill type across more than 80 sites	COD, BOD, alkalinity, pH, TDS, ammonia nitrogen, metals and trace organics	The study found that leachate characteristics vary by landfill type. MSW landfill leachates showed high ammonia nitrogen, ash related leachates showed high TDS, and all leachates had COD much greater than BOD, indicating recalcitrant organic matter.

2.3 IMPACT OF LEACHATE ON SOIL PROPERTIES

The characteristics described in Section 2.2 become environmentally significant when leachate interacts with the soil profile below and around a landfill. Soil forms the first receiving medium for leachate seepage, and the response of soil depends on texture, clay content, organic matter, pH, redox condition, permeability, moisture condition and the thickness of the unsaturated zone (Kumar & Alappat, 2005; Kanmani & Gandhimathi, 2013). Soil may retain part of the pollutant load through adsorption, ion exchange, precipitation, complexation and filtration, but the same soil can also become a contaminated medium after prolonged exposure (Kumar & Alappat, 2005; Rani et al., 2020).

Leachate interaction can change soil physicochemical properties by increasing electrical conductivity, soluble salt content, nutrient concentration, organic matter load and metal concentration. Chloride and total dissolved solids are especially important because they remain mobile and can indicate leachate migration through soil (Rani et al., 2020). High chloride concentration has been reported in Bhalswa leachate, and chloride may not be effectively attenuated by soil because of its mobile nature (Rani et al., 2020). Soil quality deterioration due to landfill leachate has also been reported across seasons and sampling locations near the Achan landfill in the temperate Himalayan region (Islam et al., 2024).

Heavy-metal accumulation is one of the most important effects of leachate-contaminated soil because metals do not degrade biologically and may remain stored in the soil matrix. Heavy-metal contamination in soil has been reported near open dumping sites due to leachate migration, with metals such as Mn, Pb, Cu and Cd showing enrichment in affected soil (Kanmani & Gandhimathi, 2013). Elevated Cr, Ni and Pb have also been reported in soil and groundwater around the Bhandewadi dumpsite in Nagpur, where the contamination pattern was linked to leachate migration from the uncontrolled dumpsite (Drall et al., 2025). These findings

show that landfill-affected soil can act as both a sink for metals and a delayed source of contamination when environmental conditions change (Kanmani & Gandhimathi, 2013; Drall et al., 2025).

The retention of metals in soil depends on adsorption on clay minerals, interaction with organic matter, precipitation, hydrolysis, co-precipitation and binding with iron and manganese oxides. However, these retention processes are not permanent because rainfall infiltration, changing leachate strength, pH variation and redox changes can remobilize retained metals and support downward migration (Kumar & Alappat, 2005; Rani et al., 2020). Therefore, landfill-affected soil should not be interpreted only as a protective barrier, because long-term pollutant loading can convert the soil into a secondary source of contamination for deeper soil layers and groundwater (Kanmani & Gandhimathi, 2013; Drall et al., 2025).

Leachate contamination can also modify the engineering behaviour of soil. Changes in Atterberg limits, compaction characteristics, shear strength and permeability have been reported in leachate-contaminated soil (Mir et al., 2025; Harun et al., 2025). Leachate exposure has been shown to affect plasticity, compaction, shear strength and permeability in landfill soil from Achan, Srinagar (Mir et al., 2025). Leachate contamination has also been found to alter the geotechnical behaviour of granitic sandy-clayey soil, including reduction in shear strength and changes in plasticity and hydraulic behaviour (Harun et al., 2025). These findings are relevant because changes in soil strength and permeability can influence contaminant movement as well as ground stability near landfill sites (Mir et al., 2025; Harun et al., 2025).

Recent literature has also shown that leachate-affected soil should be studied through soil-health and vadose-zone indicators. Long-term landfill leachate pollution can suppress soil health indicators in natural ecosystems (Kooch et al., 2024). Long-term leachate pollution can also

change pollutant distribution and microbial community structure in vadose-zone soils (Li et al., 2024).

For Bhalswa landfill, the soil impact pathway is particularly important because the site is old, receives a large municipal waste load and has strong leachate characteristics (Rani et al., 2020; CGWB, 2024). The subsurface around Bhalswa consists of fine to medium sand mixed with coarse gravel up to about 40 m below ground level, followed by older alluvium mainly consisting of clay, silt and kankar (CGWB, 2024). Sandy and gravelly layers can support contaminant movement, whereas clay-rich layers can support partial retention of metals and other pollutants (CGWB, 2024). Therefore, the present study needs to examine soil not only as a contaminated medium but also as a controlling layer that influences contaminant retention, migration and groundwater impact around Bhalswa landfill (CGWB, 2024).

Overall, the literature shows that landfill leachate can change soil properties through physicochemical alteration, salinity build-up, heavy-metal accumulation, modification of geotechnical behaviour and disturbance of soil biological functioning (Rani et al., 2020; Kanmani & Gandhimathi, 2013; Kooch et al., 2024; Mir et al., 2025). This interpretation is directly relevant to the present thesis because changes in soil behaviour can influence both the storage of contaminants within the soil profile and their eventual movement towards groundwater (CGWB, 2024; Li et al., 2024). Thus, **Section 2.3** links leachate characteristics with soil response and provides the basis for evaluating soil and groundwater behaviour around Bhalswa landfill. Table 2.2 summarizes key studies on leachate-induced changes in soil properties.

Table 2.2 Summary of Literature on Leachate-Induced Changes in Soil Properties

S. No.	Author / Year	Study area / focus	Soil parameters Studied	Major findings
1	Rani et al. (2020)	Bhalswa, Okhla and Ghazipur landfill leachate, Delhi	Chloride mobility, heavy metals, soluble salts, organic load	The study reported very high chloride concentration in Bhalswa leachate, 9853 mg/L, and stated that soil may not attenuate chlorides effectively because chlorides are highly mobile. Heavy metals may accumulate in soil through precipitation, adsorption, co-precipitation and interaction with clay and organic matter.
2	Kanmani and Gandhimathi (2013)	Open dumping site, Ariyamangalam, Tiruchirappalli	Heavy metals in soil due to leachate migration	The study assessed heavy-metal contamination in soil caused by leachate migration from an open dumping site. The paper is widely cited for showing that landfill leachate can enrich surrounding soil with metals and create a pathway for later groundwater contamination.
3	Biswas et al. (2010)	Mathkal municipal solid waste dumpsite, Kolkata	Landfilled waste characteristics, leachate, soil/subsoil percolation, groundwater impact	The study reported that leachates were characterized for Cr, Cd, Ni, Mn and Pb along with pH, COD and hardness. It noted that leachate can percolate through soil strata and pollute underlying groundwater, and that remedial measures are required to protect environmental quality around the landfill.
4	Anand et al. (2021)	Review of MSW landfill leachate and emerging contaminants	Soil physicochemical properties, biological properties, ARGs, ARB, MRGs and emerging contaminants	The review stated that released landfill leachate affects soil physicochemical and biological properties, groundwater quality, agricultural activity and human health. It also emphasized emerging contaminants such as antibiotic-resistant bacteria, antibiotic resistance genes and metal resistance genes.
5	Bai et al. (2023)	Landfill leachate-contaminated soil remediation	Organic contaminants, ammonia nitrogen, Cu, Zn, Cd, soil toxicity and plant growth	The study tested soil flushing with biosurfactant saponin and found removal efficiencies of 47.01% for organic contaminants, 90.42% for ammonia nitrogen, 29.42% for Cu, 22.55% for Zn and 17.68% for Cd. It also reported that flushing reduced plant toxicity of contaminated soil.
6	Beinabaj et al. (2023)	Tehran landfill	Heavy metals in leachate, soil and plants; landfill age effect	The study showed that soil and plants around the landfill were contaminated with heavy metals under the influence of leachate. It also found that landfill age affected the intensity and pattern of metal pollution.
7	Kooch et al. (2024)	Natural woodland and rangeland ecosystems in a semi-arid environment	Soil health indicators under long-term leachate pollution	The study investigated long-term landfill leachate pollution and found suppression of soil health indicators in natural ecosystems.
8	Mir et al. (2025)	India; Achan landfill, Srinagar	Atterberg limits, compaction characteristics, shear strength and permeability	The study examined landfill soil mixed with leachate concentrations from 0% to 20%. Leachate addition reduced Atterberg limits; with increasing leachate concentration, dry unit weight declined and optimum moisture content increased. The study also reported changes in shear strength and permeability behaviour.

S. No.	Author / Year	Study area / focus	Soil parameters Studied	Major findings
9	Harun et al. (2025)	Malaysia; granitic sandy-clayey soil contaminated with landfill leachate	Particle size distribution, pH, XRD, SEM, specific gravity, organic content, Atterberg limits, hydraulic conductivity and shear strength	Leachate contamination significantly changed soil properties. The study found reduced shear strength, reduced liquid and plastic limits, and increased permeability. It also reported that shear strength declined by up to 64.5%, indicating serious loss of soil stability.
10	Drall et al. (2025)	India; Bhandewadi dumpsite, Nagpur, Maharashtra	Heavy-metal contamination in soil and groundwater; geoaccumulation index, contamination factor, ecological risk and PCA	The study found heavy contamination in soil and groundwater near the uncontrolled dumpsite, especially around Bidgoan Road and Pardi. Elevated Cr, Ni and Pb were reported within a 2 km radius, and PCA confirmed the influence of leachate migration from the dumpsite.
11	Kooch et al. (2024)	Iran; natural woodland and rangeland ecosystems near landfill sites	Soil bulk density, organic matter, porosity, aggregate stability, particulate organic carbon, nitrogen, nutrients, microbial biomass, enzymes and soil biota	Long-term landfill leachate pollution increased soil bulk density and reduced soil organic matter, porosity, aggregate stability, available nutrients, microbial biomass and enzyme activities. Soil cadmium and lead also showed strong spatial variation under landfill influence.
12	Li et al. (2024)	China; informal landfill site in Northern China	Vadose-zone soil pollutants, depth-wise contamination, heavy metals, soil characteristics and microbial community structure	The study examined long-term leachate pollution across soil depths. It found that leachate movement from surface to deeper soil layers changed microbial diversity and that nitrate, humic fractions and toxic metal fractions were major drivers of microbial community change. Soil characteristics alone explained 57.9% of bacterial community variation.
13	Kooch et al. (2024)	Iran; natural woodland and rangeland ecosystems near landfill sites	Soil bulk density, organic matter, porosity, aggregate stability, particulate organic carbon, nitrogen, nutrients, microbial biomass, enzymes and soil biota	Long-term landfill leachate pollution increased soil bulk density and reduced soil organic matter, porosity, aggregate stability, available nutrients, microbial biomass and enzyme activities. Soil cadmium and lead also showed strong spatial variation under landfill influence.
14	Islam et al. (2024)	Achan landfill, Kashmir, India	Soil quality across seasons and dumping-site zones	Leachate seepage affected soil quality, with greater deterioration near the centre of the dumping site.

*Antibiotic Resistance Genes (ARGs), Antibiotic-Resistant Bacteria (ARB), Metal Resistance Genes (MRGs),

2.4 GROUNDWATER CONTAMINATION DUE TO LANDFILL LEACHATE

Groundwater contamination due to landfill leachate occurs when pollutants move from the landfill body through soil and the unsaturated zone into the aquifer. This movement depends on leachate strength, soil permeability, depth to groundwater, rainfall infiltration, hydraulic gradient, and the presence or absence of landfill lining and leachate collection systems (Kumar & Alappat, 2005; Sharma et al., 2020).

In landfill-affected areas, groundwater quality deterioration is commonly reflected through high electrical conductivity, total dissolved solids, chloride, nitrate, sulphate, hardness, ammoniacal nitrogen, heavy metals, and bacteriological contamination (Alam et al., 2020; Sharma et al., 2020). Parameters such as EC, TDS and chloride are especially useful because they indicate dissolved ionic load and possible leachate migration into groundwater (Rani et al., 2020; CGWB, 2024).

For Bhalswa landfill, groundwater contamination is particularly important because the site is located in an alluvial setting with shallow groundwater and nearby residential areas. Groundwater around Bhalswa has been reported as brackish, with high EC, nitrate, fluoride, chloride, heavy metals and bacteriological contamination in several samples (CGWB, 2024). Earlier studies near Bhalswa also reported groundwater quality deterioration due to the influence of non-engineered landfill conditions and leachate migration (Srivastava & Ramanathan, 2008; Ahamad et al., 2019).

Heavy metals are a major concern in landfill-affected groundwater because they are persistent and may pose long-term environmental and health risks. Studies near landfill sites have reported contamination by Fe, Mn, Pb, Cd, Cr, Ni, As, Cu and Zn, depending on waste composition and local hydrogeological conditions (Boateng et al., 2019; Ahamad et al., 2019). This is relevant to Bhalswa because site-specific reports have also indicated heavy metal concerns in groundwater and leachate samples (CGWB, 2024).

The spatial and seasonal pattern of groundwater contamination is also important. Contamination often decreases with increasing distance from the landfill, while monsoon infiltration can influence leachate movement and contaminant mobilization (Alam et al., 2020; Sharma et al., 2020). Therefore, groundwater assessment around Bhalswa should consider distance from landfill, groundwater flow direction, seasonal variation and index-based classification such as WQI and HPI (Sharma et al., 2020; CGWB, 2024).

Overall, landfill leachate can deteriorate groundwater through salinity increase, ionic enrichment, heavy metal contamination and microbial pollution. For the present thesis, Section 2.4 helps establish groundwater as the affected resource after leachate movement through soil, and supports the need for hydrochemical assessment, spatial analysis and predictive modelling around Bhalswa landfill. Table 2.3 summarizes key studies on groundwater contamination due to landfill leachate and their relevance to the present study around Bhalswa landfill.

Table 2.3 Summary of Literature on Groundwater Contamination due to Landfill Leachate

S. No.	Author / Year	Study area	Parameters / methods used	Key findings
1	CGWB (2024)	Bhalswa landfill site, NCT Delhi	EC, pH, fluoride, chloride, nitrate, heavy metals, bacteriological parameters, pre- and post-monsoon assessment	Groundwater around Bhalswa showed high EC, salinity concerns, fluoride, nitrate, chloride and bacteriological contamination. Shallow aquifers close to the landfill were reported to be influenced by leachate migration.
2	Ahamad et al. (2019)	Bhalswa landfill, Yamuna flood plain, New Delhi	LPI, WQI, major ions, heavy metals, health risk assessment, pre- and post-monsoon sampling	Groundwater near Bhalswa showed elevated EC, major cations, anions and heavy metals. WQI was used to classify drinking suitability, and health risk was assessed through ingestion exposure.
3	Srivastava and Ramanathan (2008)	Vicinity of Bhalswa landfill, Delhi	Hydrogeochemistry, nitrate, fluoride, heavy metals, graphical and multivariate methods	Groundwater near Bhalswa showed hydrogeochemical alteration, with nitrate, fluoride and heavy metal concerns linked to landfill leaching and local hydrogeological conditions.
4	Alam et al. (2020)	Ghazipur landfill, New Delhi	pH, EC, TDS, alkalinity, chloride, sulphate, nitrate, BOD_5, COD, iron, coliforms and hardness	Groundwater quality near Ghazipur was affected by leachate percolation, and contamination decreased with increasing distance from the landfill.
5	Angmo and Shah (2021)	Okhla, Bhalswa and Ghazipur landfill sites, Delhi	Chloride, nitrate, sulphate, ammonium, phenol, iron, zinc, COD and groundwater quality assessment	Groundwater quality around Delhi landfill sites was affected by leachate percolation, especially where uncontrolled dumping and lack of lining were present.
6	Sharma et al. (2020)	Solan, Mandi, Sundernagar and Baddi dumpsites, Himachal Pradesh	LPI, WQI, HPI, PCA, HCA, downstream distance-based groundwater characterization	Leachate pollution affected groundwater quality, and water quality improved with increasing distance from dumpsites. PCA and HCA helped identify pollution zones.
7	Boateng et al. (2019)	Oti landfill site, Kumasi, Ghana	Pb, Fe, Cd, Cr, Zn, Cu, HPI, health risk assessment and multivariate analysis	Mean Pb, Fe, Cd and Cr exceeded WHO drinking water limits, and high metal levels were linked to leachate percolation from the landfill.
8	Kumar and Alappat (2005)	Landfill leachate pollution assessment	LPI, organic sub-index, inorganic sub-index and heavy metal sub-index	The study established that landfill leachate pollution potential can be quantified through LPI and sub-indices, including parameters relevant to groundwater contamination.

2.5 APPLICATION OF MACHINE LEARNING IN GROUNDWATER QUALITY PREDICTION AND SOIL HEAVY METAL PREDICTION

Machine learning has become an important tool in environmental quality assessment because groundwater and soil contamination are controlled by several interrelated parameters. Conventional statistical methods are useful for identifying general trends, but machine learning models can capture non-linear relationships among physicochemical parameters, hydrogeological variables, spatial indicators and contamination indices. In landfill-affected environments, this is particularly useful because contaminant movement is influenced by leachate strength, soil behaviour, groundwater flow, distance from the pollution source and seasonal variation. Therefore, machine learning can support groundwater quality prediction and soil heavy metal prediction by improving classification, spatial interpretation and decision-making based on measured environmental datasets.

a. Groundwater Quality Prediction

Machine learning has been widely applied for groundwater quality prediction using indices such as Water Quality Index, Groundwater Quality Index, entropy-weighted water quality index, nitrate pollution index and GIS-based vulnerability indices. These studies generally use input parameters such as pH, electrical conductivity, total dissolved solids, major cations, major anions, nitrate, chloride, sulphate, hardness and groundwater level to predict overall groundwater suitability or contamination risk. Models such as Random Forest, XGBoost, Artificial Neural Network, Support Vector Machine, Decision Tree, Gaussian Process Regression and ensemble learning methods have shown strong performance in predicting groundwater quality classes and contamination indices (Elmotawakil et al., 2025; El-Said Salem et al., 2025; Abba et al., 2024; Yadav et al., 2024).

Recent studies show that tree-based and ensemble models often perform well because they can handle complex and non-linear relationships among water quality parameters. XGBoost showed

very high prediction accuracy in water quality classification when compared with LSTM, Decision Tree, SVM, Random Forest and ANN models (Elmotawakil et al., 2025). Similarly, GIS-based GALDIT and composite vulnerability index prediction using XGBoost showed strong performance for groundwater vulnerability assessment in areas affected by seawater intrusion and aquifer sensitivity (El-Said Salem et al., 2025). These studies indicate that machine learning can be useful not only for predicting water quality, but also for identifying vulnerable groundwater zones.

Several studies have also used machine learning to predict specific groundwater contamination indicators. Gaussian Process Regression performed best for predicting nitrate concentration in groundwater when compared with Elman Neural Network and Random Forest models (Abba et al., 2024). Random Forest also performed strongly for entropy-weighted water quality index prediction using hydrochemical parameters such as pH, TDS, fluoride, potassium, sodium, magnesium, calcium, nitrate, sulphate, chloride and total hardness (Yadav et al., 2024). These findings show that machine learning can support both index-based classification and parameter-specific prediction in groundwater studies.

Ensemble and hybrid models have further improved groundwater quality prediction by combining the strengths of different algorithms. A quadratic stacking ensemble learning approach achieved high training and testing accuracy for Groundwater Quality Index prediction using parameters such as nitrate, electrical conductivity, potassium, calcium, magnesium, sulphate, chloride, sodium, pH and TDS (Karimi et al., 2024). ANN combined with optimization methods also performed well in predicting groundwater potential zones by using groundwater levels, hydraulic conductivity, landsat imagery, digital elevation model and related spatial variables (Abd-Elmoboud et al., 2024). These approaches are important because groundwater quality is influenced by both hydrochemical and spatial factors.

In the present study, machine learning is applied to groundwater quality prediction using 488 groundwater samples and input parameters such as pH, EC, chloride, magnesium, calcium, potassium and total hardness. The Water Quality Index is used as the prediction target, and different models such as K-Nearest Neighbors, Support Vector Machine, Decision Tree, Random Forest, XGBoost and Artificial Neural Network are compared. The results indicate that ANN and SVM showed high predictive performance, with validation showing ANN as the superior model. This supports the use of machine learning as a predictive tool for groundwater quality assessment around landfill-affected areas, especially where multiple hydrochemical parameters jointly influence water quality classification.

The literature shows that groundwater quality prediction has moved from simple index calculation towards data-driven modelling and spatial classification. For the Bhalswa landfill study, this is relevant because groundwater contamination is not controlled by one parameter alone; rather, it depends on the combined influence of salinity, ions, hardness, leachate migration and local hydrogeological conditions. Machine learning therefore provides a useful supporting approach for identifying groundwater quality patterns and improving environmental assessment around the landfill site. The Table 2.4 summarizes comparative analysis of WQI for ML models.

Table 2.4 Summary of Comparative analysis of WQI for ML MODELS

S. No	ML Technique	Best Model	Prediction Index	Dataset	Input parameters	Accuracy	References
1	LSTM, XGBoost, DT, SVM, RF, ANN	LSTM, XGBoost	Water quality classification (WQC)	Multiple Monitoring Stations; 61,542 instances	Temperature, Specific conductance, Salinity, Dissolved oxygen, Depth, pH, and Turbidity	XGBoost; Accuracy = 99.83%, LSTM; R^2= 0.99	Elmotawakkil et al., 2025
2	XGBoost, SVM, Generalized Linear Model (GLM)	XGBoost	Geographic Information System (GIS)-based GALDIT index, Conditioning Vulnerability Index (CVI)	371 wells and satellite imagery	Groundwater Occurrence, Aquifer Hydraulic Conductivity, Groundwater Level above sea level, Distance from Shoreline, Impact of Existing Seawater Intrusion, and Aquifer Thickness.	R^2 =96.22%	El-Said Salem et al.,2025
3	Elman Neural Network (ELNN), Random Forest (RF), and Gaussian Process Regression (GPR)	Gaussian Process Regression (GPR)	Wolffman Index (WI) for predicting nitrate concentration in groundwater	72 groundwater samples	Chlorine, sodium, calcium, magnesium, electrical conductivity (EC), potassium, bicarbonate, nitrate, water salinity	99.99%	Abba et al., 2024
4	Deep Neural Network (DNN), XGBoost, Randon Forest	Random Forest	Entropy-weighted water quality index (EWQI)	977 wells	pH, TDS, Fluoride, Potassium, Sodium, Magnesium, Calcium, Nitrate, Sulphate, Chloride, Total Hardness	RF: R^2 =98%(Training), R^2 = 89.7% (Testing) with low RMSE(Testing) =52.62; XGB: R^2 =98%(Training), R^2 = 91.2% (Testing); DNN: R^2 =91%(Training), R^2 = 92.2% (Testing)	A. Yadav et al., 2024
5	AdaBoost classifier (ADA), Quadratic discriminant analysis (QDA), and stacking ensemble learning (SEL)	Quadratic stacking ensemble learning (SEL)	Groundwater Quality Index (GWQI)	250 wells	NO^-_3, EC, K^+, Co^{2+}, Mg^{2+}, SO^{2-}_4, Cl, Na, pH, TDS	Training Overall Accuracy= 95%, Testing Overall Accuracy =91%	Karimi et al., 2024
6	ANN, Mountain gazelle optimization (MGO).	ANN with MGO	Groundwater Potential zones (GWPZs)	8 wells	Groundwater Levels, Hydraulic Conductivity, Landsat image, Digital elevation model (DEM)	AUC = 90%	Abd-Elmaboud et al., 2024

S. No	ML Technique	Best Model	Prediction Index	Dataset	Input parameters	Accuracy	References
7	SVM, Naïve Bayes, DT	DT	WAWQI	Naramada River,India water samples	pH, EC, Dissolved Oxygen (DO), Biochemical Oxygen Demand (BOD)	Accuracy = 98.50%	Radhakrishnan and Pillai, 2020
8	Convolutional Neural Networks (CNN), KNN, Naive Bayes (NB), Multi-Layer Perceptron (MLP).	CNN	Water Quality Index	379 Groundwater samples	pH, EC, Groundwater level, Latitude, Longitude, TDS, CO_3, HCO_3, Cl, F, NO_3, SO_4, Na, K, Ca, Mg, TH	R^2 =0.99	Huang et al., 2025
9	K-Nearest Neighbors, Support Vector Machine, Decision Tree, Random Forest, XGBoost, and Artificial Neural Network	ANN and SVM, But Validation Proves ANN Superior	Water Quality Index (WA-WQI), Taylor Diagram technique, Spatial Distribution of WQI	488 Groundwater samples	pH, EC, Cl, Mg, Ca, K, TH	R^2 = 0.99 for ANN and SVM; Mean Absolute Percentage Error = 3.27(ANN), 10.36(SVM)	This Study

b. Soil Heavy metal Prediction

Machine learning is increasingly used for soil heavy metal prediction because metal concentration in soil is influenced by several interacting physical, chemical, spatial and anthropogenic factors. Soil properties such as pH, organic carbon, clay content, nutrients, electrical conductivity and background metal concentration are commonly used as input variables, while land use, topography, distance-related variables and pollution-source indicators help capture the spatial behaviour of contamination (Zhang et al., 2023; Zhang et al., 2020; Wang et al., 2020). These models are useful because heavy metal distribution is rarely uniform, and relationships between soil properties and metal accumulation are often non-linear (Wang et al., 2020).

Recent studies show that tree-based and ensemble models perform well in soil heavy metal prediction because they can handle mixed environmental, spatial and chemical variables. XGBoost has shown strong performance for cadmium prediction using soil properties such as EC, pH, $CaCO_3$, silt and phosphorus, while machine learning models have also been applied to predict soil adsorption of Cd, Cr, Cu, Pb, Ni and Zn using large experimental datasets (Keçeci et al., 2024; Yang et al., 2021). For contaminated-site assessment, machine learning is also useful for identifying possible sources and spatial patterns of heavy metal pollution. Source-identification studies have used soil properties, land-use indicators, topographic variables and pollution-source information to distinguish natural and anthropogenic controls on heavy metal distribution (Jia et al., 2019; Wang et al., 2020).

In the present study, soil heavy metal prediction is carried out for cadmium, chromium and lead using physicochemical parameters and associated metal concentrations as input variables. Models such as RF, XGBoost, SVM, ANN, Linear Regression, Logistic Regression, RF-GA and NN-ANFIS are compared to identify the best prediction approach. The use of hybrid RF-

GA is relevant because optimized models can improve prediction accuracy when soil heavy metal behaviour is influenced by multiple interacting parameters. The following Table 2.5 summarizes key machine learning studies related to soil heavy metal prediction and their relevance to the present study.

Table 2.5 Summary of key machine learning studies related to soil heavy metal prediction

S.No	Author	Model used	Variables	Outcomes	Remarks
1	Zhang et al., 2023	ANN Decision Trees SVM Naive Bayes ANFIS KNN Random Forest XGBoost	Covariates co-occurring with pollutants Soil pH Transportation-related variables Industrial activity variables Topographic variables (e.g. elevation, slope)	Soil contaminant concentrations or levels (e.g. heavy metals like Cd, Cr, Hg, Pb, Ni, Cu, As, Zn, and organic pollutants like TPAHs)	The paper proposes a structured guideline for using ML to predict soil contamination, which includes guidance on variable selection, model evaluation, and model
2	Paes et al., 2022	Random Forest, Cubist, Support Vector Machine, K-Nearest Neighbor, Linear Regression	X-ray fluorescence (XRF) data Morphometric data derived from a digital elevation model (DEM) Euclidean distance to the main drainage channel (EUDC) Soil texture data (coarse sand, fine sand, silt, clay) Magnetic susceptibility (MS) data	Ba, Cr, Zn, Pb, Fe, Al	The Cubist machine learning model was the best-performing for predicting the contents of most potentially toxic elements in the soil.
3	Zhang et al., 2020	Random Forest ANN SVM	Land use types: built-up area (BA), area under construction (CA), cropland (CL), forestland (FL), and waterbody (WB) Soil properties: Fe2O3 and CaO, Organic Carbon, Total N, P, S	The concentrations of As, Zn, Pb, Hg, Ni, Cu, Cr, and Cd in the soil	SVM had the highest accuracy. The paper aims to identify the relationship between urbanization and heavy metal accumulation in urban soils, and to use machine learning models to predict the spatial distribution of heavy metals based on variables related to urban expansion
4	Bhagat et al., 2021	XGBoost ANN SVM	Concentration of heavy metals like Hg, Fe, Cd, As, Ti, U, Th, etc.	Concentration of Pb in sediments of two Australian bays	XGBoost had the highest accuracy.
5	Cao et al., 2020	Deep Composite Model (DCM)	Longitude, latitude, altitude, functional area and eight different soil heavy metal contents.	Prediction of content of heavy metals in soil.	The deep composite model (DCM) proposed in the paper achieved good prediction accuracy for soil heavy metal content and outperformed other models like RBFNN, RRBFNN, and ARBFNN.

S.No	Author	Model used	Variables	Outcomes	Remarks
		combining an RBFNN			
6	Bazoobandi et al., 2019	ANFIS ANN MLP	Clay Organic carbon (O.C.) pH Phosphorus (P) Total nitrogen (T.N.)	Cadmium and Lead concentration in soil	The ANN model outperformed the ANFIS and MLR models in predicting cadmium and lead content in the soil, with lower RMSE values.
7	Wang et al., 2020	Random Forest Land Use Regression	Population indicators Physical and chemical properties of the soil Sewage irrigation indicators Atmospheric deposition indicators Industrial pollution indicators Household waste indicators Transportation indicators Land use indicators	Soil heavy metal(loid) concentrations (Pb, Cd, Cr, As, Hg, and Zn)	Random Forest model outperforms the land use regression model and the results can be used for soil resource and environmental quality management.
8	This Study	RF, XGBoost, Support Vector Machines (SVM), Artificial Neural Networks (ANN), Logistic Regression (LR), RF-GA(Genetic Algorithm) model	pH, EC, TDS, chlorides, alkalinity, total hardness, and Heavy Metals: Mg, Al, Mn, Fe, Zn, Cu, Cd, Cr, and Pb.	Heavy Metals: Cadmium (Cd), Chromium (Cr) and Lead (Pb)	Hybrid RF-GA model consistently outperformed the other five ML models, particularly with larger training datasets, demonstrating high specificity and rapid convergence to lower cost function values.

2.6 SYNTHESIS OF LITERATURE REVIEW AND RELEVANCE TO THE PRESENT STUDY

The literature reviewed in this chapter provides the scientific basis for examining landfill-leachate, induced changes in soil and groundwater behaviour around Bhalswa landfill. Previous studies establish that landfill leachate is a major contaminant source containing organic matter, dissolved salts, nitrogenous compounds, chlorides, heavy metals and microbial contaminants, while groundwater-focused studies show its influence on electrical conductivity, salinity, nitrate, fluoride, heavy metals and bacteriological quality (Kumar & Alappat, 2005; Rani et al., 2020; Alam et al., 2020; Sharma et al., 2020; CGWB, 2024).

Soil-focused studies further indicate that leachate interaction can alter soil physicochemical properties, promote heavy metal accumulation, affect soil health indicators and influence geotechnical behaviour (Kanmani & Gandhimathi, 2013; Kooch et al., 2024; Mir et al., 2025). These findings support a source–pathway–receptor understanding of the present study, where leachate is treated as the contamination source, soil as the receiving and controlling medium, and groundwater as the affected resource.

This integrated interpretation is particularly relevant for Bhalswa landfill because the site is old, receives a large municipal waste load and is located in an alluvial setting where sand, gravel, silt, clay and kankar can influence contaminant movement and retention (CGWB, 2024). Machine learning studies also support the use of data-driven models for groundwater quality prediction and soil heavy metal prediction when contamination is controlled by multiple interacting parameters (Jia et al., 2019; Yang et al., 2021; Aju et al., 2024). Thus, the reviewed literature connects leachate characteristics, soil response, groundwater contamination and predictive modelling, and provides the foundation for the methodology adopted in this thesis.

CHAPTER 3: PHYSICOCHEMICAL, HEAVY METAL AND GEOTECHNICAL ASSESSMENT OF LANDFILL LEACHATE-IMPACTED SOIL

3.1 INTRODUCTION

This chapter presents the physicochemical, heavy metal and geotechnical assessment of landfill leachate-impacted soil around the Bhalswa landfill. Landfill leachate is examined as the primary contaminant source, while soil samples collected from different depths and distances are analysed to assess the extent of contamination and its possible influence on soil behaviour.

Landfill leachate is a complex liquid generated through the interaction of waste, moisture, rainfall and decomposition processes. Its composition depends on waste characteristics, landfill age, climatic conditions, drainage pattern and stage of waste degradation. It commonly contains dissolved solids, organic matter, chloride, sulphate, nitrogenous compounds, alkalinity, hardness and potentially toxic metals, which makes leachate characterization necessary for understanding its environmental impact (Kumar & Alappat, 2005). Recent studies have also reported the presence of conventional and emerging contaminants in landfill leachate, including heavy metals, metalloids, organic pollutants and microplastics (Qian et al., 2024; Wisitthammasri et al., 2024).

In this chapter, leachate samples are analysed to determine their pollutant strength and contamination potential. Parameters such as pH, electrical conductivity, total dissolved solids, salinity, alkalinity, hardness, chloride, sulphate, nitrate, phosphate, calcium and magnesium are considered to understand the chemical nature of the leachate. Heavy metals such as Cd, Cr, Pb, Ni, Cu, Zn, Fe, Mn and As are also examined because of their persistence, toxicity and potential mobility through soil under favourable environmental conditions (Hussein et al., 2021; Saha et al., 2023). The leachate pollution index is further used to provide an integrated assessment of leachate

contamination level. Figure 3.1 illustrates the field view of MSW accumulation and visible leachate seepage at Bhalswa landfill, Delhi.

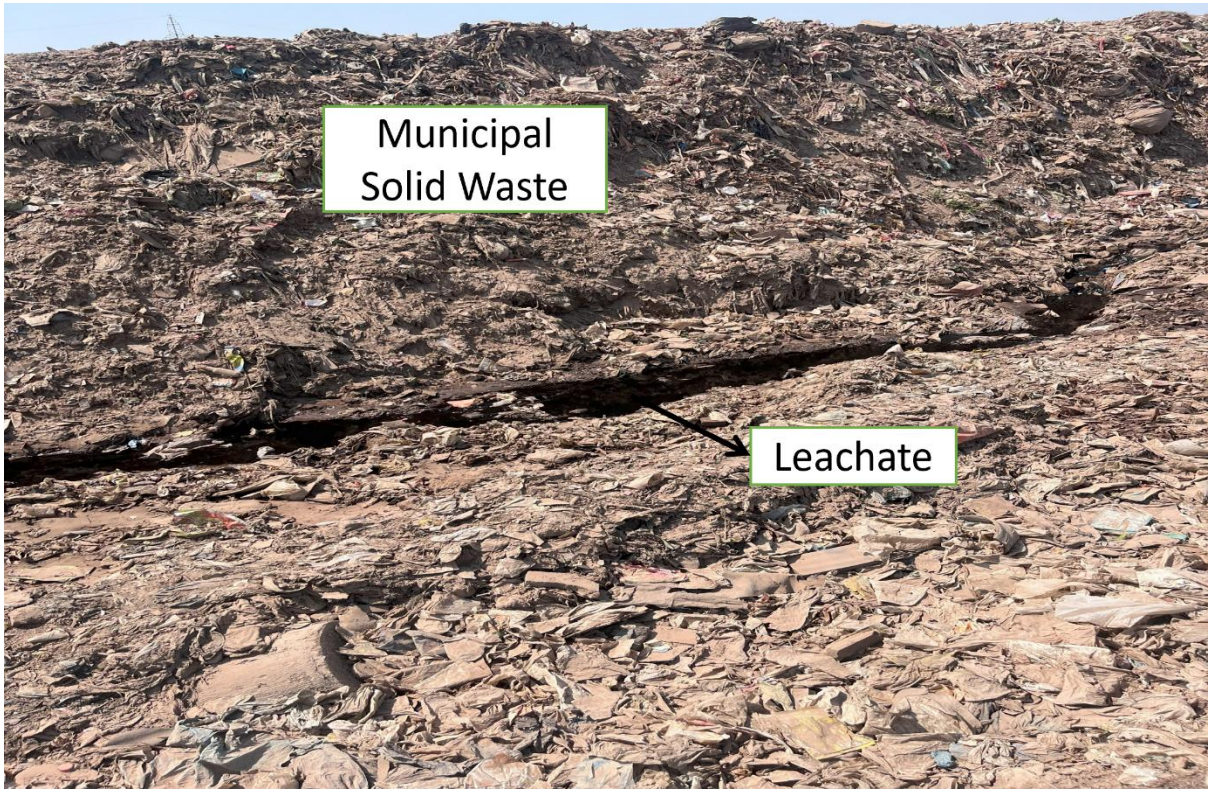


Figure 3.1 Field view of MSW accumulation and visible leachate seepage at Bhalswa landfill, Delhi.

Soil assessment is included because soil acts as the first receiving medium after leachate release from the landfill body. During leachate–soil interaction, soil properties may change in terms of pH, EC, TDS, salinity, chloride, alkalinity, nutrient content and heavy metal concentration. These changes depend on soil texture, mineral composition, organic matter content, drainage conditions and duration of exposure (Kanmani & Gandhimathi, 2013; Islam et al., 2024). Therefore, soil samples are analysed to evaluate the spatial and vertical variation of contamination around the landfill site. The heavy metal component is important because metals are persistent and non-biodegradable. These may be accumulated in soil through adsorption, precipitation, complexation and interaction with clay minerals or organic matter, but may also become mobile with changing pH, redox and moisture conditions. Pollution indices such as geo-accumulation index and

enrichment factor are used to distinguish between natural background influence, anthropogenic enrichment and stronger contamination levels.

In addition to chemical assessment, selected geotechnical properties of soil are examined to understand whether leachate contamination has influenced soil behaviour. Parameters such as Index and compaction characteristics, permeability, cohesion, angle of internal friction and moisture response are considered because leachate interaction may alter soil structure and engineering behaviour (Harun et al., 2013; Khodary et al., 2018). SEM–EDX analysis is also used to support the interpretation of soil microstructure and elemental composition.

Overall, this chapter provides an integrated assessment of leachate quality, soil contamination, heavy metal distribution, pollution indices and geotechnical behaviour around the Bhalswa landfill. The findings help identify contamination intensity, affected zones and possible pathways of leachate influence on soil quality. These results also provide a basis for interpreting groundwater quality variations discussed in the **Chapter 4** and for supporting suitable mitigation measures for landfill-affected areas.

3.2 MATERIALS AND METHODS

3.2.1 STUDY AREA DESCRIPTION

The present study was conducted in the vicinity of Bhalswa landfill, Delhi, located at approximately 28.7395° N latitude and 77.1573° E longitude. Bhalswa landfill is one of the major landfill sites of Delhi and is situated in Model Town tehsil of North District. The landfill occupies an area of about 52 acres, equivalent to 21.06 ha, and receives nearly 2,500 metric tonnes of municipal solid waste per day (CPCB 2024, Bharadwaj et al.,2025a). The filling of waste at the site was commissioned in 1994, and the height of the dump is reported to be approximately 60–

65 m above the surrounding ground level. The elevation of the area ranges from about 206 to 221 m above mean sea level (CPCB 2024).

The climate of the region is semi-arid, with hot summers and cold winters. The normal annual rainfall of NCT Delhi is about 611.8 mm, of which nearly 81% occurs during the monsoon months of July, August and September. The landfill receives mixed municipal solid waste, including biodegradable waste, plastics, inert materials, construction waste and other non-biodegradable fractions (Bharadwaj et al.,2025b). Due to the unlined and non-engineered nature of the site, leachate generation and its movement through the subsurface are important environmental concerns (CGWB 2024).

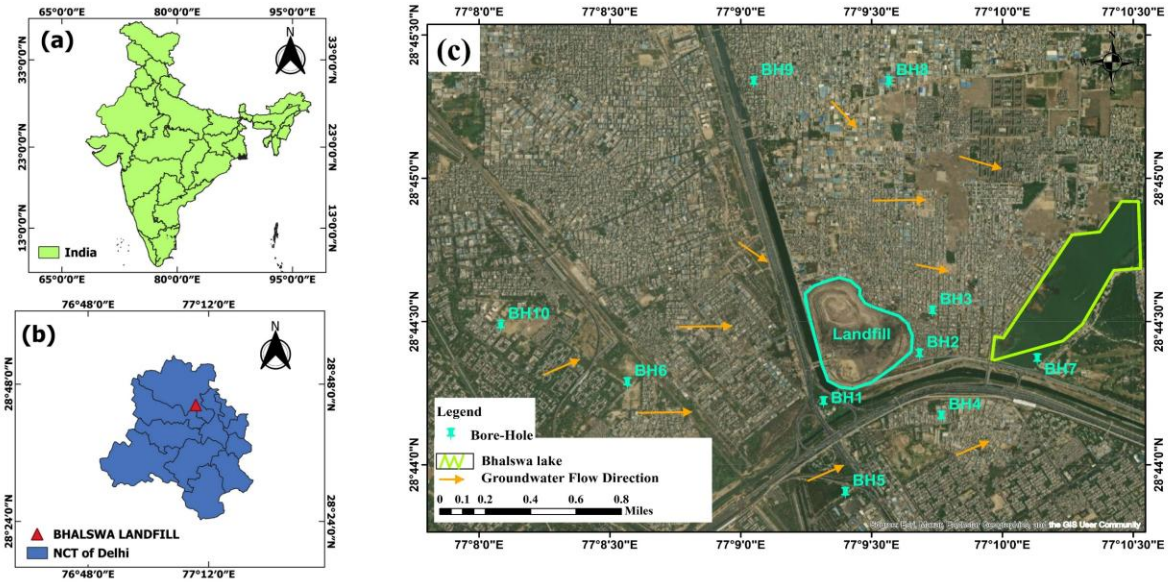


Figure 3.2 Location map showing soil sampling points in the vicinity of Bhalswa landfill, Delhi.

Geologically, the study area is mainly underlain by Quaternary alluvial deposits. The alluvium is about 200 m thick, below which weathered and fractured quartzite is present. Based on lithological and geophysical observations, the upper subsurface layer consists mainly of fine to medium sand, locally known as Yamuna sand, mixed with coarse gravel up to a depth of about 40 m below ground level. This layer is underlain by older alluvium consisting predominantly of clay,

silt and kankar. The sandy and gravelly nature of the upper layer may facilitate the infiltration and movement of landfill leachate into the subsurface (CGWB 2024). Figure 3.2 illustrates the location map showing soil sampling points in the vicinity of Bhalswa landfill, Delhi.

3.2.2 SAMPLE COLLECTION AND PRESERVATION

Leachate Sample Collection

A total of five fresh leachate samples were collected from designated leachate ponds located at the base of the waste heaps at the Bhalswa landfill site, Delhi, in December 2023, where leachate accumulates through gravity drainage (Figure 3.3).



Figure 3.3 Leachate Collection Point at Bhalswa Landfill

The samples were collected in clean, sterile high-density polyethylene (HDPE) containers, which were pre-rinsed with dilute nitric acid and deionised water to minimise potential contamination. Each container was properly sealed and labelled with the sample code, location and date of collection. Immediately after sampling, the leachate samples were transported to the analytical laboratory in insulated containers maintained at approximately 4 °C using frozen ice packs to

reduce biological and chemical changes during transit. The analysis of leachate samples was carried out within 24 hours of collection. In the laboratory, the preserved samples were filtered using Whatman filter paper, wherever required, to remove flaky and coarse suspended impurities before analysis. The analysis was performed as per the relevant IS 3025 and APHA (American Public Health Association standard methods.

Soil Sample Collection

A total of 50 soil samples were collected in January 2024, through borehole drilling in the vicinity of the Bhalswa landfill dumpsite, New Delhi, India. Ten borehole locations were selected to provide adequate spatial coverage around the landfill, and five samples were collected from each borehole at 3 m depth intervals, i.e., 3 m, 6 m, 9 m, 12 m and 15 m (Figure 3.4).

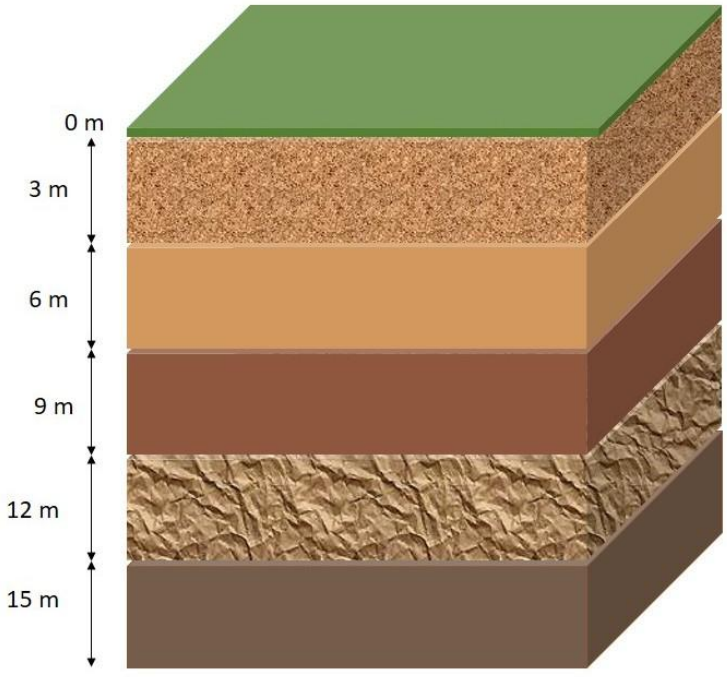


Figure 3.4 Schematic representation of soil sampling

The sampling design was planned to assess the influence of landfill leachate contamination in both horizontal and vertical directions. A benchmark point was established at a suitable location near the Bhalswa landfill, from which the distances of the borehole locations ranged from 306.81 m to 2012.93 m (Figure 3.5). The boreholes were designated as BH-1 to BH-10, as shown in Figure 3.2. After collection, the soil samples were immediately placed in airtight zip-lock polyethylene bags to prevent external contamination and minimise moisture loss. Each sample was properly labelled with borehole number, depth and sampling location. The collected samples were then transported to the laboratory for further processing. In the laboratory, the samples were oven-dried for 24 hours at 80°C before subsequent physicochemical, heavy metal and geotechnical analyses. The soil sampling points in the vicinity of Bhalswa landfill, Delhi, India are shown in Figure 3.6.

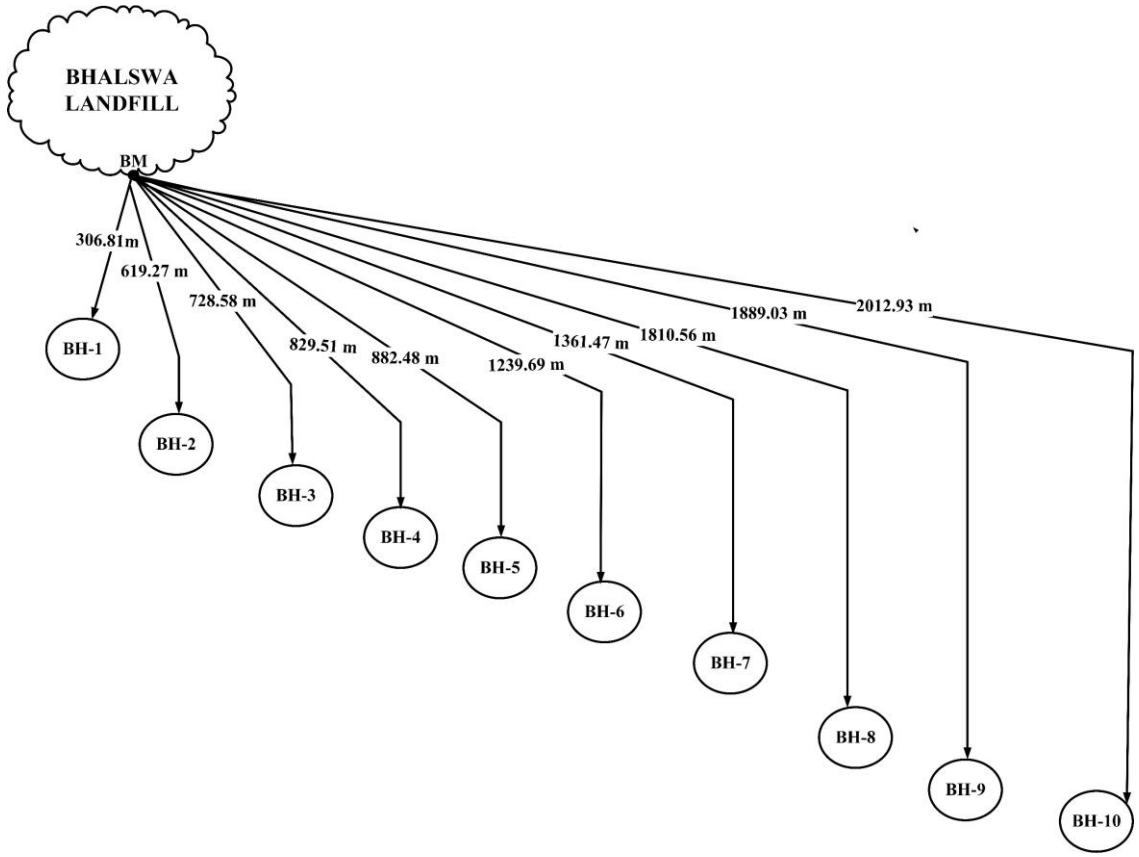


Figure 3.5 Schematic representation of boreholes distances from Bhalswa landfill, Delhi, India



Figure 3.6 Soil sampling points in the vicinity of Bhalswa landfill, Delhi, India

3.2.3 LABORATORY ANALYSIS

Assessment of leachate samples

The collected five leachate samples were analysed to assess their physicochemical characteristics, organic pollution load, ionic composition and heavy metal contamination potential. The parameters determined for leachate included pH, electrical conductivity, total dissolved solids, total suspended solids, alkalinity, total hardness, calcium, magnesium, chloride, sulphate, nitrate, phosphate, chemical oxygen demand and selected heavy metals such as Cu, Ni, Pb, Cr, Fe, Cd and Al. These parameters were selected to evaluate the overall pollution strength of the leachate and

its potential influence on the surrounding soil environment. All analyses were carried out using the relevant IS 3025 and APHA standard methods, and the parameter-wise method codes used for leachate analysis are presented in Table 3.1.

Table 3.1 Methods Used for Leachate Sample Analysis

S. No.	Parameter	Method/Standard Used
1	pH	IS 3025 Part 11 / APHA 4500-H(B)
2	Chemical Oxygen Demand (COD)	IS 3025 Part 58 / APHA 5220(B)
3	Total Suspended Solids (TSS)	IS 3025 Part 17 / APHA 2540(D)
4	Total Dissolved Solids (TDS)	IS 3025 Part 16 / APHA 2540(C)
5	Electrical Conductivity (EC)	IS 3025 Part 14 / APHA 2510(B)
6	Phosphate	IS 3025 Part 31 / APHA 4500-P(D)
7	Nitrate	IS 3025 Part 34 / APHA 4500-NO_3(B)
8	Sulphate	IS 3025 Part 24 / APHA 4500-SO_4(E)
9	Total Alkalinity	IS 3025 Part 23 /APHA 2320 B
10	Total Hardness	IS 3025 Part 21 / APHA 2340 C
11	Calcium	IS 3025 Part 40 / APHA 3500-Ca B
12	Chloride	IS 3025 Part 32 / APHA 4500-Cl^- B
13	Copper (Cu)	IS 3025 Part 42 / APHA 3500-Cu B
14	Lead (Pb)	IS 3025 Part 47 / APHA 3500-Pb B
15	Nickel (Ni)	IS 3025 Part 54 / APHA 3500-Ni B
16	Chromium (Cr)	IS 3025 Part 52 / APHA 3500-Cr B
17	Cadmium (Cd)	IS 3025 Part 41 / APHA 3500-Cd D
18	Iron (Fe)	IS 3025 Part 53/ APHA 3500-Fe B
19	Aluminium (Al)	IS 3025 Part 55/ APHA 3500-Al B

Physicochemical and Heavy Metal Characterization of Soil samples

The collected soil samples (Figure 3.7) were analysed to evaluate their physicochemical characteristics and metal contamination status in the vicinity of the Bhalswa landfill. The physicochemical parameters included pH, electrical conductivity, total dissolved solids, salinity, resistivity, alkalinity, chloride, calcium and total hardness. These parameters were selected to assess the ionic strength, salinity condition and chemical alteration of soil due to possible leachate interaction. The metal analysis included Zn, Cd, Cu, As, Ag, Co, Mg, Ni, Fe, Mo, Al, Pb, Mn, Cr and K. For metal analysis, the soil samples were subjected to acid digestion, followed by

instrumental determination using an MP-AES (Microwave Plasma Atomic Emission Spectroscopy). The concentrations of metals/elements and solution-derived chemical parameters were reported in mg/kg on a dry soil weight. The instruments and methods used for soil sample analysis are presented in Table 3.2.

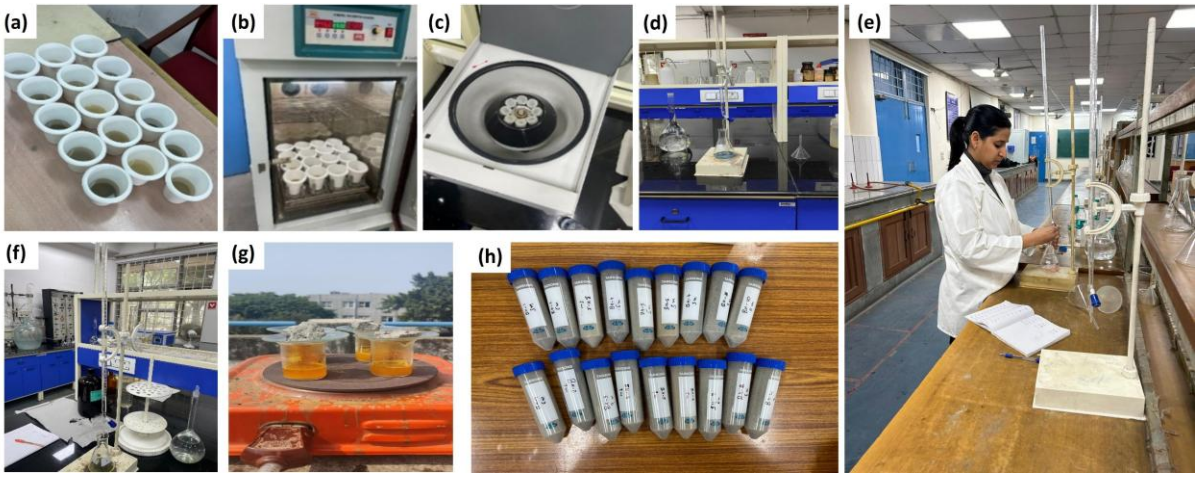


Figure 3.7 Laboratory testing (a) Saturation of Soil Samples (b) Orbital shaker with soil samples in flasks (c) Centrifuge machine with soil samples (d) Titration for total hardness (e) Performing test in lab (f) Volumetric titration for Chloride (g) Acid Digestion of Soil Sample (h) Prepared samples for testing

Table 3.2 Instruments/Methods Used for Physicochemical and Heavy Metal Analysis of Soil Samples

S. No.	Parameter	Unit	Instrument/Method Used
1	pH value	—	Multiparameter Electrode (IS 2720 Part 26:1987)
2	Electrical conductivity	µS/cm	Multiparameter Electrode, IS 14767:2000 / Soil Testing Manual, Ministry of Agriculture, Govt. of India
3	Total dissolved solids / Total soluble solids	mg/kg	TDS meter, IS 2720 Part 21:1977
4	Salinity	psu	Multiparameter Electrode
5	Resistivity	kΩ-cm	Multiparameter Electrode
6	Alkalinity	mg/kg	Acid-base titration
7	Chloride	mg/kg	Argentometric titration
8	Total hardness as $CaCO_3$	mg/kg	EDTA titration
9	Calcium hardness as $CaCO_3$	mg/kg	EDTA titration
10	Magnesium hardness as $CaCO_3$	mg/kg	EDTA titration
11	Calcium as Ca	mg/kg	USEPA 7000B:2007 / Soil Testing Manual, Ministry of Agriculture, Govt. of India
12	Zn, Cd, Cu, As, Ag, Co, Mg, Ni, Fe, Mo, Al, Pb, Mn, Cr and K	mg/kg	Acid digestion followed by MP-AES (Soil Testing Manual, Ministry of Agriculture, Govt. of India)
13	Potassium as K	mg/kg	Acid digestion followed by MP-AES (IS 9497:1980)

Geotechnical Characterization of Soil

The collected soil samples were subjected to geotechnical characterization to evaluate the influence of landfill leachate contamination on soil behaviour. The analysis included determination of natural water content, particle size distribution (Sieve analysis and hydrometer analysis), consistency limits, specific gravity, compaction characteristics, permeability (variable-head) and shear strength parameters. The particle size distribution was evaluated in terms of gravel, sand, silt + clay and clay fractions, while the consistency behaviour was assessed using liquid limit, plastic limit and plasticity index. Based on the particle size distribution and plasticity characteristics, the soil samples were classified as per IS 1498:1970. Compaction behaviour was assessed through optimum moisture content and maximum dry density, and percentage changes in OMC and MDD were used to examine variations in compaction response among the samples. The hydraulic behaviour was evaluated through the coefficient of permeability, expressed as $K \times 10^{+6}$ m/s. Shear strength behaviour was assessed using cohesion and angle of internal friction. These parameters were selected to understand changes in soil strength, compressibility, permeability and engineering behaviour due to possible leachate interaction.

3.2.4 DATA ANALYSIS AND POLLUTION ASSESSMENT METHODS

Descriptive Statistical Analysis

Descriptive statistics of heavy metal concentrations were used to provide a preliminary understanding of the overall distribution, variability, and concentration range of metals in the soil samples. Parameters such as mean, minimum, maximum, standard deviation, and coefficient of variation were calculated for each metal to identify the general abundance pattern and degree of variation across the 50 soil samples. The mean values indicate the average concentration level of each metal, while minimum and maximum values show the concentration range observed in the

study area. Standard deviation and coefficient of variation help assess whether the metals are uniformly distributed or show high spatial/depth-related variability. Metals with high variability may indicate localized contamination, anthropogenic input, or uneven distribution due to landfill influence, soil properties, or leachate movement. Thus, descriptive statistics provide the basic foundation for further analysis such as correlation, regression, enrichment factor, and geo-accumulation index assessment.

Correlation and Regression Analysis

Correlation and regression analysis were carried out to evaluate the relationship of heavy metal concentrations with sampling depth and landfill proximity. A Spearman correlation heatmap was prepared to visually examine the overall association among heavy metals, sampling depth, and landfill proximity. The heatmap represents the strength and direction of monotonic relationships between variables, where positive correlation values indicate that two variables tend to increase together, while negative values indicate an inverse relationship. In this study, the inclusion of sampling depth and distance from the landfill helped assess whether heavy metal concentrations showed any systematic vertical or horizontal variation across the 50 soil samples. The heatmap also helped identify metals showing similar distribution behaviour, which may indicate common sources, comparable mobility, or similar geochemical control in the soil environment.

Pollution Index Assessment

Geo-accumulation Index (Igeo) and Enrichment Factor (EF) were used to assess the contamination intensity and possible anthropogenic enrichment of heavy metals in soil samples. Igeo was calculated by comparing the measured concentration of each metal with its respective background value, using a correction factor to account for natural lithogenic variation. The obtained Igeo

values were classified into different contamination categories ranging from uncontaminated to extremely contaminated. EF was calculated to evaluate whether the observed metal concentrations were mainly derived from natural crustal sources or influenced by anthropogenic inputs. In the present study, Fe was used as the reference element for EF calculation, and metals with EF values greater than 1.5 were considered to indicate possible anthropogenic enrichment.

Leachate Pollution Index (LPI) was used to assess the overall pollution potential of landfill leachate based on selected physicochemical and chemical parameters. For each leachate parameter, the sub-index value was obtained using the standard rating curves proposed by Kumar and Alappat, 2003 and the corresponding parameter weight was used to calculate the final LPI value. The LPI provides a single numerical value representing the cumulative pollution strength of leachate, thereby helping to understand its possible contribution to soil and groundwater contamination around the landfill site. Together, Igeo, EF, and LPI were used to link heavy metal contamination in soil with landfill-related pollution intensity and potential leachate influence.

3.3 LANDFILL LEACHATE CHARACTERIZATION

Landfill leachate characterization was carried out to assess the pollution strength and contaminant potential of leachate generated from the Bhalswa landfill site. Five leachate samples were analysed for physicochemical parameters, organic pollution load, major ions and selected heavy metals. The results were compared with reference values prescribed under BIS IS 10500:2012 and the Solid Waste Management Rules, 2016, wherever applicable. Since landfill leachate is a highly concentrated wastewater, these reference values were used to understand the relative contamination intensity and environmental risk associated with leachate migration into the surrounding soil system.

3.3.1 PHYSICOCHEMICAL AND HEAVY METAL CHARACTERISTICS OF LEACHATE

The analysed leachate samples (Figures 3.8) showed alkaline behaviour, with pH values within the reference range of 5.5–9.0. However, most of the other physicochemical parameters indicated high pollution strength. Electrical conductivity and total dissolved solids were considerably elevated, suggesting a high concentration of dissolved ionic constituents in the leachate. Chloride concentration was higher than the reference limit in all samples, indicating strong salt enrichment and the possible influence of decomposed municipal solid waste. The leachate also showed elevated concentrations of nitrate, phosphate, total hardness, alkalinity and chemical oxygen demand, reflecting the presence of dissolved salts, nutrient load and organic contamination.

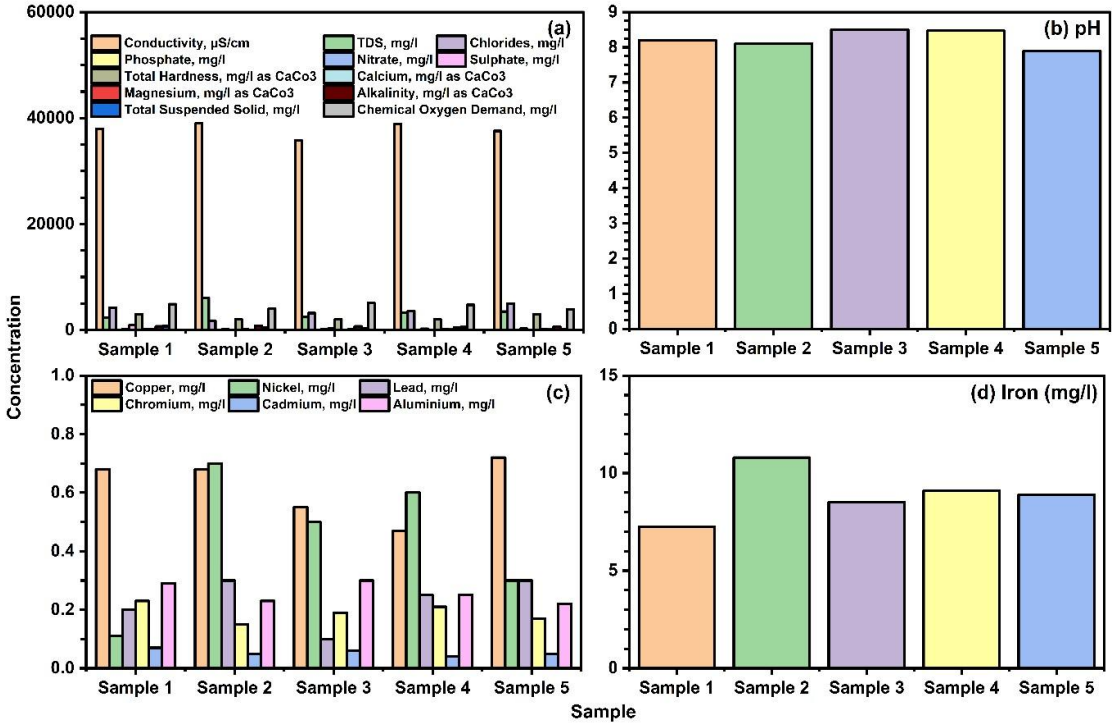


Figure 3.8 Leachate sample quality results

The chemical oxygen demand values were substantially higher than the reference value, indicating a strong organic pollution load and the presence of oxidizable organic matter in the leachate. Total

suspended solids were also above the reference limit in all samples, suggesting the presence of suspended impurities and particulate matter. Among the ionic constituents, chloride, nitrate and total hardness showed marked enrichment, confirming the chemically aggressive nature of the leachate.

Table 3.3 Physicochemical and Heavy Metal Characteristics of Leachate Samples Collected from Bhalswa Landfill

Parameters	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Reference Value BIS IS 10500:2012 / SWM Rules 2016
Chlorides, mg/l	4260	1775	3195	3550	4970	1000
Phosphate, mg/l	18.9	11.74	20.83	22.37	18.83	5
Nitrate, mg/l	195	100	220	234	254	45
Sulphate, mg/l	908	42	262	82	66	400
Total Hardness, mg/l as $CaCo_3$	3000	2000	2000	2000	3000	600
Calcium, mg/l as $CaCo_3$	160	120	80	162	180	200
Magnesium, mg/l as CaCo3	113	79	98	86	94	100
Alkalinity, mg/l as $CaCo_3$	700	800	700	500	600	600
Total Suspended Solid, mg/l	712	470	325	560	210	100
Chemical Oxygen Demand, mg/l	4890	4,000	5,100	4,750	3,900	250
Copper, mg/l	0.68	0.68	0.55	0.47	0.72	1.5
Nickel, mg/l	0.11	0.7	0.5	0.6	0.3	0.02
Lead, mg/l	0.2	0.3	0.1	0.25	0.3	0.01
4.5mium, mg/l	0.23	0.15	0.19	0.21	0.17	0.05
Iron, mg/l	7.26	10.8	8.5	9.1	8.9	1
Cadmium, mg/l	0.07	0.05	0.06	0.04	0.05	0.003
Aluminium, mg/l	0.29	0.23	0.3	0.25	0.22	0.2
pH	8.2	8.1	8.5	8.47	7.9	5.5–9.0
Conductivity, µS/cm	38000	39000	35800	38864	37560	3000
TDS, mg/l	2340	6110	2440	3270	3490	500

Heavy metal analysis showed the presence of Cu, Ni, Pb, Cr, Fe, Cd and Al in the leachate samples. Among these, Ni, Pb, Cr, Fe, Cd and Al exceeded their respective reference values, indicating significant heavy metal contamination potential. Iron was present at relatively high concentrations in all samples, while lead, cadmium and chromium also showed values above the reference limits. Copper remained within the adopted reference value in the analysed samples. The overall results indicate that Bhalswa landfill leachate is characterized by high ionic strength, organic load and heavy metal enrichment, making it a potential source of soil and groundwater contamination. Table 3.3 illustrates physicochemical and heavy metal characteristics of Bhalswa landfill leachate with reference values.

3.3.2 LEACHATE POLLUTION INDEX ASSESSMENT

The Leachate Pollution Index (LPI) was calculated to evaluate the overall pollution potential of the leachate samples collected from the Bhalswa landfill. The calculation was carried out using the original pollutant weights and sub-index rating curves proposed by Kumar and Alappat (2005) as given in Equation (3.1). Since all the 18 parameters included in the original LPI framework were not available in the present study, the index was calculated using the available LPI parameters, namely pH, total dissolved solids, chemical oxygen demand, chloride, total chromium, lead, nickel, copper and iron. Other measured parameters such as phosphate, nitrate, sulphate, calcium, magnesium, alkalinity, total hardness, total suspended solids, electrical conductivity, cadmium and aluminium were not included in the LPI calculation because they are not part of the original LPI parameter set. The calculated LPI values for the five leachate samples ranged from 14.89 to 16.71, with an average value of 15.93. The lowest LPI value was observed in Sample 2, while the highest value was observed in Sample 5. The limited variation among the five samples indicates that the leachate collected from different leachate accumulation points within the landfill

showed broadly similar pollution characteristics. The LPI values confirm that the leachate possesses considerable pollution potential and may act as an important contaminant source for the surrounding soil environment.

$$LPI = \frac{\sum_{i=1}^{n}(W_i \times P_i)}{\sum_{i=1}^{n} W_i} \quad (3.1)$$

Where,

Wi = Weight assigned to the ith pollutant variable

Pi = Sub-index score of the ith pollutant

n = Number of pollutants considered in the calculation

Among the parameters considered for LPI calculation, chemical oxygen demand and chloride contributed strongly to the index values. The high contribution of COD indicates the presence of a significant organic pollution load, while the contribution of chloride reflects the dissolved salt content and mobility potential of the leachate. Chloride is particularly important because it is generally conservative and mobile in the subsurface, and therefore can be used as an indicator of leachate migration. The contribution of total dissolved solids further supports the presence of a high dissolved ionic load in the leachate samples.

The heavy metal component of the LPI was represented by chromium, lead, nickel, copper and iron. Although the sub-index contribution of heavy metals was lower than that of COD and chloride, their presence remains environmentally significant because metals are persistent and may accumulate in soil over time. The pH values of the leachate samples were within the alkaline range and contributed comparatively less to the overall LPI. Therefore, the LPI assessment suggests that the pollution potential of Bhalswa landfill leachate is mainly controlled by organic load, dissolved salts and selected heavy metals. Overall, the average LPI value of 15.93 indicates that the leachate generated from the Bhalswa landfill has measurable contamination potential (Table 3.3). The

Bhalswa landfill is an old and stabilized landfill, which resulted in moderate LPI. The leachate generated from the landfill has the potential to pollute soil and groundwater at a higher level. When considered along with the elevated values of COD, chloride, TDS and selected heavy metals, the LPI result supports the interpretation that landfill leachate may influence the physicochemical and heavy metal characteristics of nearby soil. Hence, the LPI assessment provides an integrated basis for linking leachate quality with the observed contamination trends in leachate-impacted soil samples.

Table 3.4 Leachate pollution Index

Leachate Sample	LPI
Sample 1	16.18
Sample 2	14.54
Sample 3	15.17
Sample 4	15.53
Sample 5	16.62
Mean LPI	**15.93**

The leachate quality was further compared with the applicable reference values prescribed under the Solid Waste Management Rules, 2016 and BIS IS 10500:2012, wherever parameter-specific limits were available. The comparison showed that several parameters, including chloride, COD, TSS, TDS, nitrate, total hardness, EC and heavy metals such as nickel, lead, chromium, iron, cadmium and aluminium, exceeded the adopted reference values, indicating the chemically strong and contaminated nature of the leachate. Although BIS IS 10500:2012 is primarily a drinking-water standard and not a leachate discharge standard, it was used as a reference guideline for parameters not covered under the treated leachate standards. Overall, the exceedance of multiple parameters supports the LPI results and confirms that Bhalswa landfill leachate has considerable pollution potential.

3.4 PHYSICOCHEMICAL AND HEAVY METAL ASSESSMENT OF LEACHATE-IMPACTED SOIL

The physicochemical and heavy metal assessment of leachate-impacted soil was carried out using 50 soil samples collected from ten borehole locations around the Bhalswa landfill at depths of 3 m, 6 m, 9 m, 12 m and 15 m. The borehole locations covered a distance range of 306.81 m to 2012.93 m from the selected benchmark point near the landfill. The analysis included physicochemical parameters such as pH, electrical conductivity, total dissolved solids, salinity, resistivity, alkalinity, chloride, calcium and total hardness, along with metals including Zn, Cd, Cu, As, Ag, Co, Mg, Ni, Fe, Mo, Al, Pb, Mn, Cr and K. The results were used to examine the influence of landfill leachate on soil chemistry and metal accumulation with respect to depth and distance from the landfill.

3.4.1 PHYSICOCHEMICAL AND HEAVY METAL CHARACTERISTICS OF SOIL

The physicochemical characteristics of the soil samples showed noticeable variation across the borehole locations and depths. The pH ranged from 6.30 to 8.22, indicating slightly acidic to alkaline subsurface conditions. The lowest pH was recorded at BH-7 at 6 m, while the highest pH was observed at BH-3 at 3 m and BH-6 at 3 m. Electrical conductivity varied from 81.17 to 713.7 µS/cm, TDS from 40.27 to 350.2 mg/L, salinity from 0.092 to 0.396 psu, and resistivity from 1.401 to 12.33 kΩ-cm. The measurement of pH and electrical conductivity in all 50 soil samples, ranging from BH-1 to BH-10, is presented in Figure 3.9. The highest EC, TDS and salinity values, along with the lowest resistivity, were recorded at BH-6 at 9 m, indicating localized ionic enrichment.

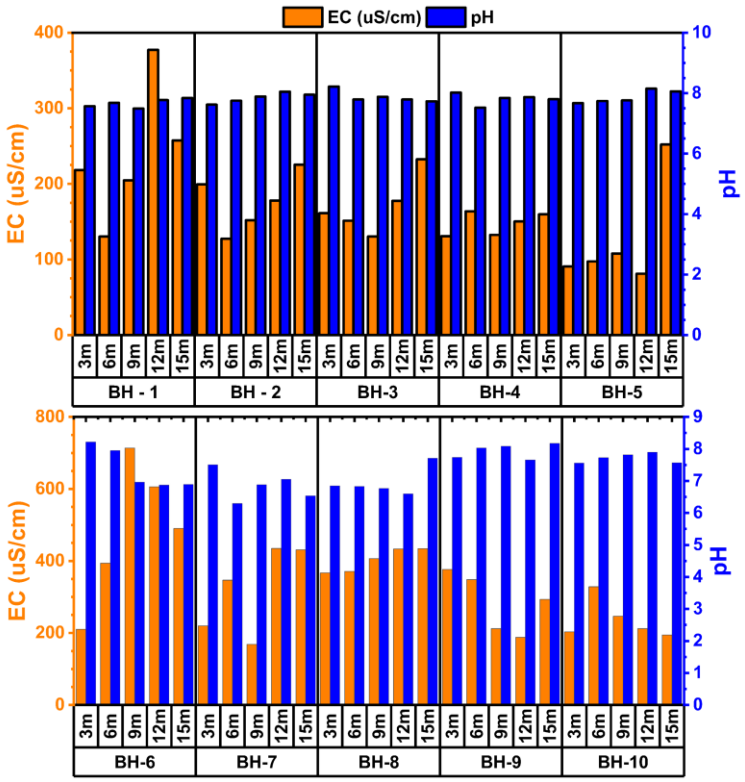


Figure 3.9 Depth-wise assessment of EC and pH in soil samples

Alkalinity ranged from 5000 to 19000 mg/kg, chloride from 998.4 to 7987.2 mg/kg, calcium from 1575 to 10500 mg/kg, and total hardness from 4500 to 20500 mg/kg. The highest alkalinity was observed at BH-4 at 3 m, while the highest total hardness was recorded at BH-9 at 15 m. The measurement of alkalinity, Cl, Ca and TH assessment in soil samples, is presented in Figure 3.10. The heavy metal concentrations were primarily compared with Indian natural background soil values reported by Kuhad et al. (1989) and Gowd et al. (2010). Where Indian natural background values were not available, Indian soil limits reported by Awasthi (2000) and WHO/FAO reference values (WHO/FAO, 2007) were used (Table 3.4). Figure 3.11 illustrates the variation in concentration for magnesium and potassium.

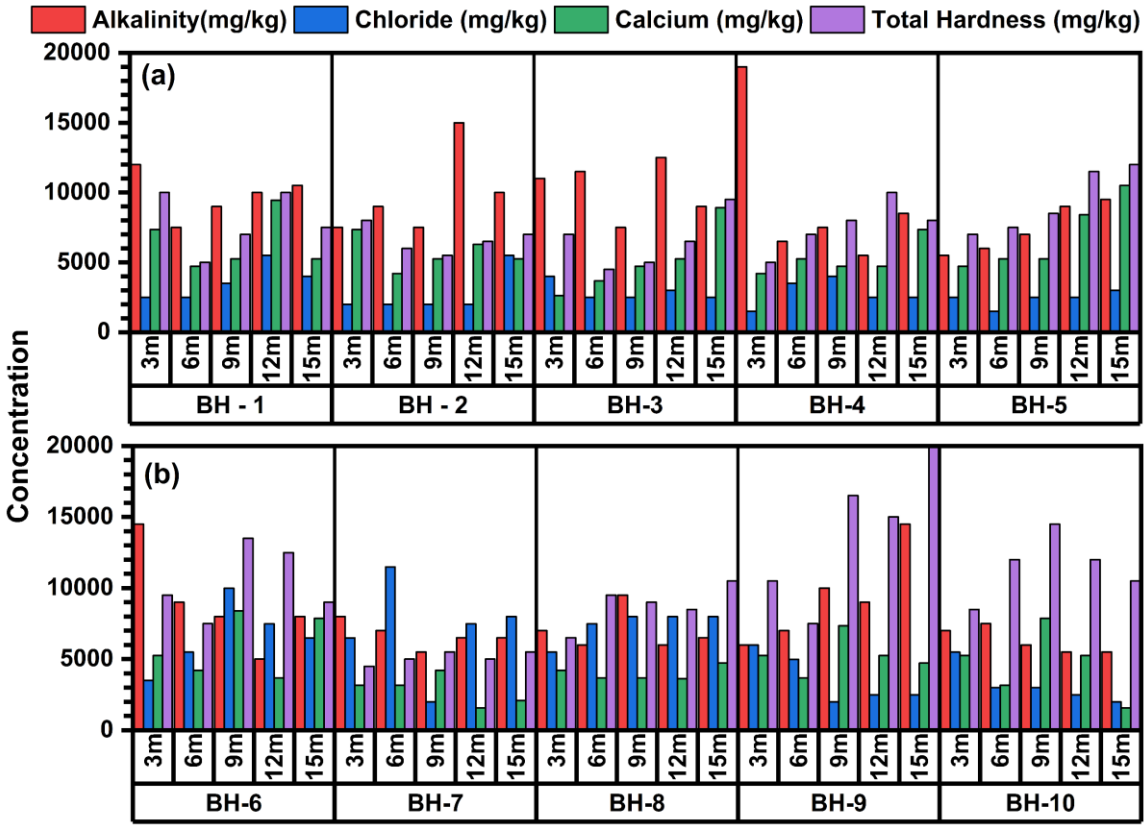


Figure 3.10 Depth-wise assessment of Alkalinity, Cl, Ca and TH in soil samples

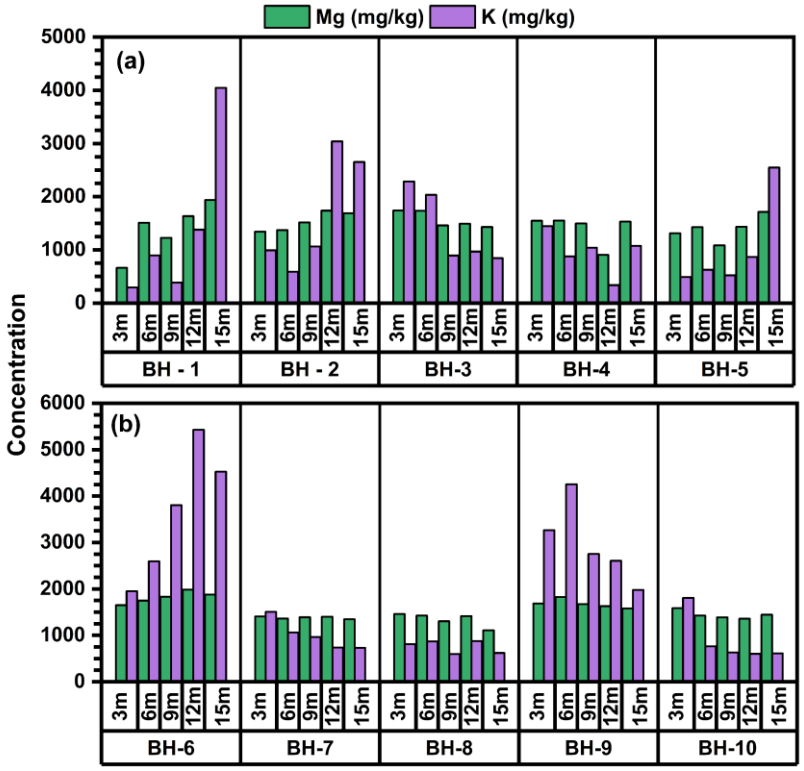


Figure 3.11 Depth-wise assessment of Mg and K in soil samples

Manganese (Figure3.12) ranged from 45 to 510 mg/kg and exceeded the Indian natural background value of 209 mg/kg in 19 out of 50 samples, with the maximum value recorded at BH-6 at 12 m. Arsenic was compared with the WHO/FAO reference value of 30 mg/kg because Indian natural background and Indian soil limit values were not reported. The As concentration ranged from 1 to 6 mg/kg, remaining below the WHO/FAO reference value in all samples (Table 3.4).

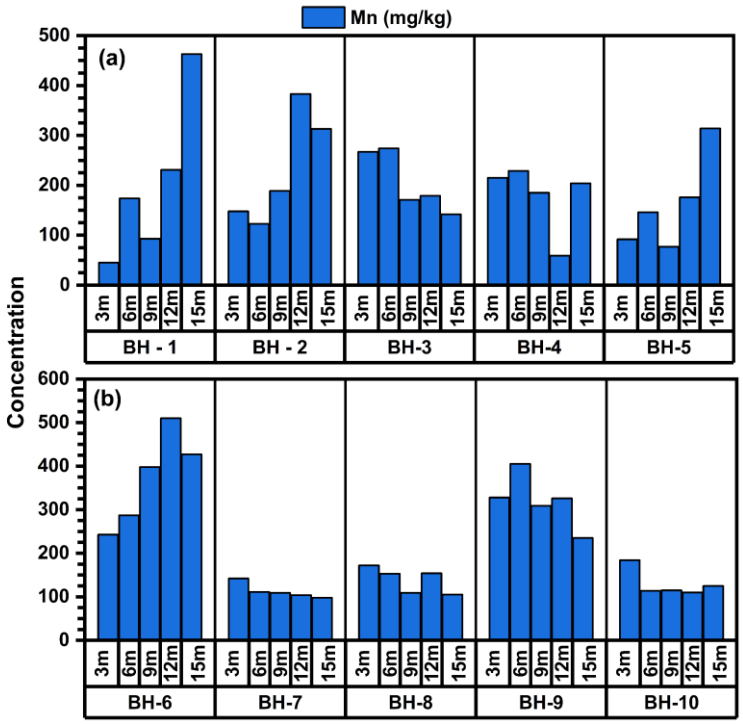


Figure 3.12 Depth-wise assessment of Mn in soil samples

Copper ranged from 2 to 43 mg/kg, which was below the Indian natural background value of 56.5 mg/kg in all samples. Chromium ranged from 4 to 56 mg/kg, remaining below the Indian natural background value of 114 mg/kg. The variation in concentration of As, Cu, Cr, and Ag is depicted in Figure 3.13 for all boreholes at each depth. Iron (Figure 3.14) ranged from 3195 to 33854 mg/kg; only BH-6 at 12 m exceeded the Indian natural background value of 32015 mg/kg. Whereas, nickel remained below the Indian natural background value of 27.7 mg/kg, with a maximum observed value of 19 mg/kg. Cobalt also remained below the Indian natural background value of 15.2 mg/kg (Kuhad et al., 1989 & Gowd et al., 2010).

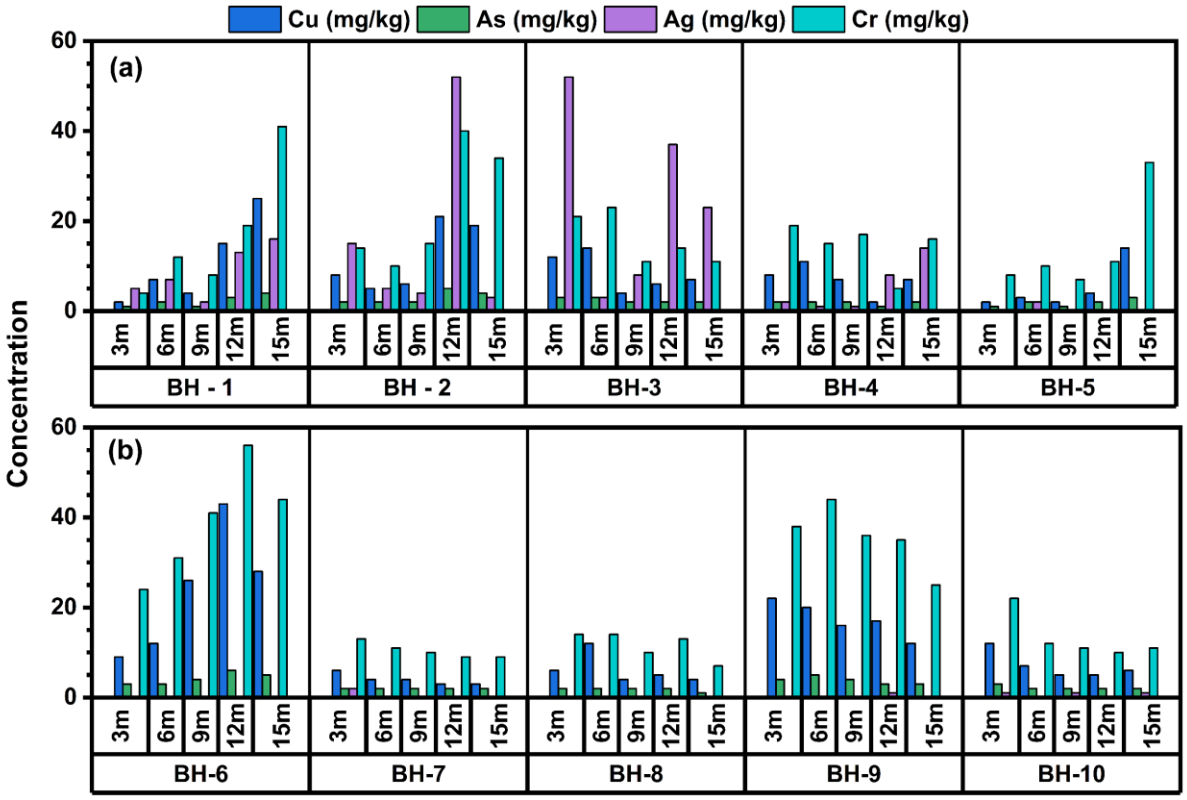


Figure 3.13 Depth-wise assessment of Cu, As, Ag and Cr in soil samples

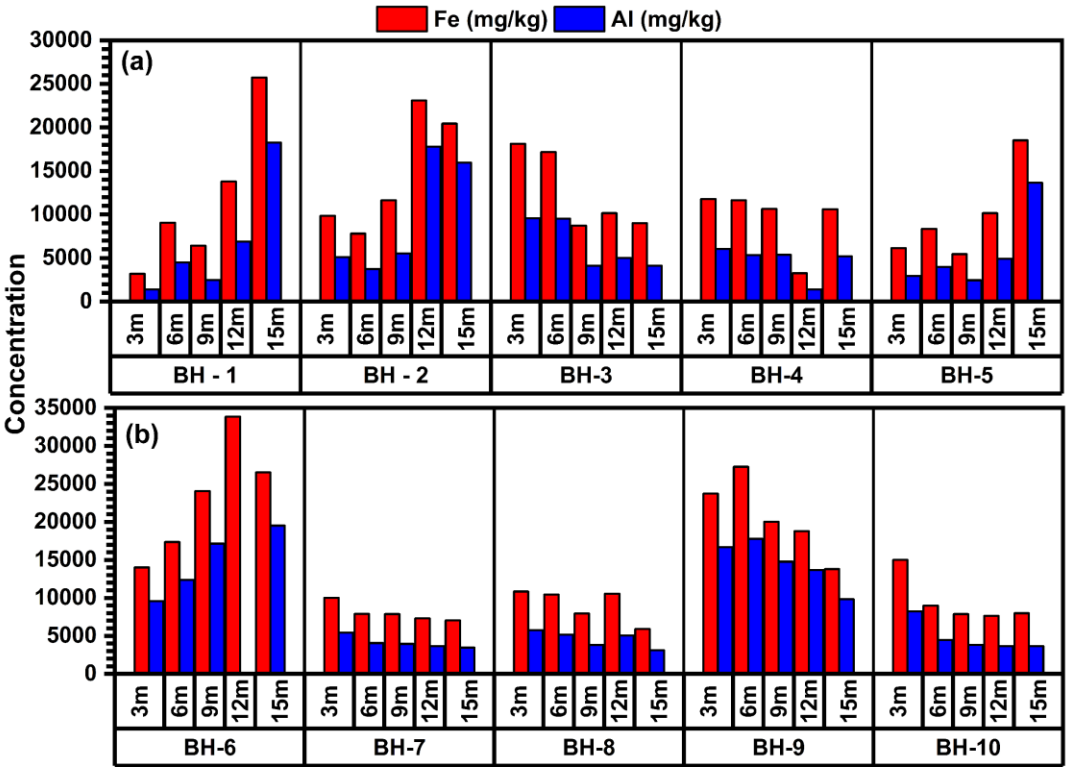


Figure 3.14 Depth-wise assessment of Fe and Al in soil samples

Zinc (Figure 3.15) ranged from 10 to 102 mg/kg and exceeded the Indian natural background value of 22.1 mg/kg in 43 out of 50 samples, with the maximum value recorded at BH-6 at 12 m. Lead(Figure 3.15) ranged from 6 to 94 mg/kg and exceeded the Indian natural background value of 13.1 mg/kg in 30 out of 50 samples; the highest Pb concentration was recorded at BH-5 at 9 m. Cadmium was compared with the Indian soil limit of 3–6 mg/kg because its Indian natural background value was not reported. Cd (Figure 3.15) ranged from 1 to 9 mg/kg; values above the lower limit of 3 mg/kg were observed at BH-2 at 3 m, BH-4 at 15 m, BH-5 at 3 m, BH-5 at 9 m, BH-6 at 3 m, BH-6 at 6 m, BH-6 at 9 m, BH-6 at 12 m, BH-6 at 15 m and BH-7 at 3 m. The upper limit of 6 mg/kg was exceeded at BH-2 at 3 m, BH-5 at 9 m and BH-6 at 3 m, with the maximum Cd concentration of 9 mg/kg recorded at BH-5 at 9 m.

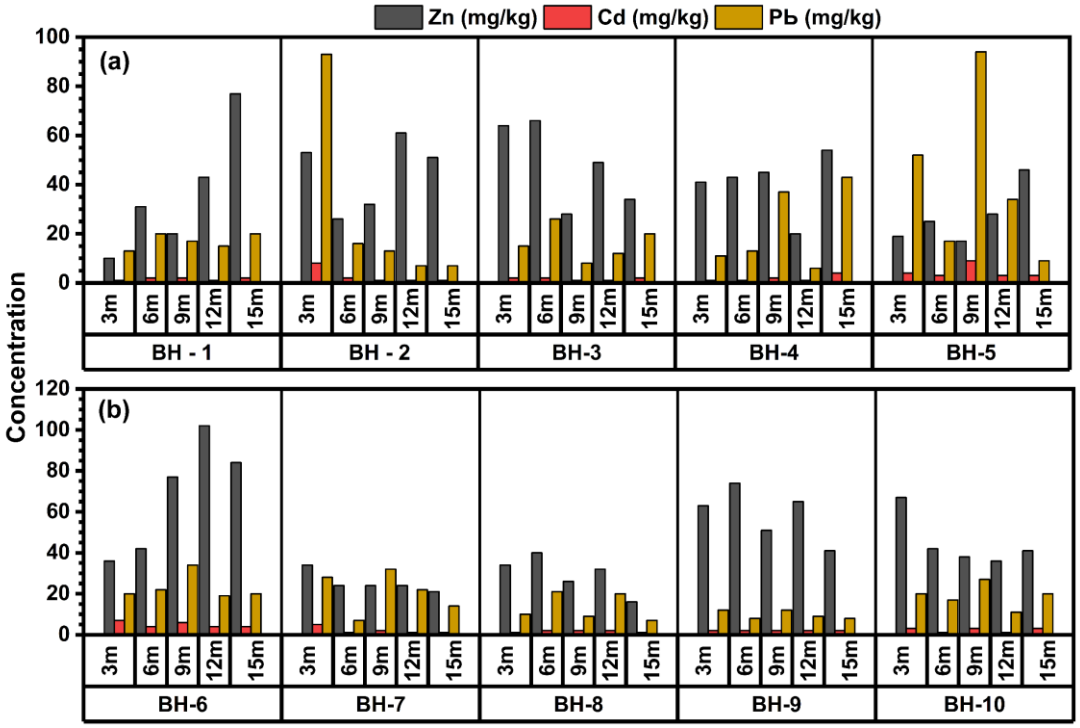


Figure 3.15 Depth-wise assessment of Zn, Cd and Pb in soil samples

The exceedance pattern indicates that Zn and Pb were the dominant metal contaminants, followed by Mn and Cd. Zinc showed the highest frequency of exceedance, crossing the Indian natural background value in 43 out of 50 samples, while Pb exceeded its background value in 30 samples.

Mn exceeded the background value in 19 samples, whereas Cd exceeded the lower Indian soil limit in 10 samples and the upper limit in 3 samples. Therefore, the overall contamination significance may be expressed as **Zn > Pb > Mn > Cd > Fe > Cu > Cr > Ni > As > Co**. Table 3.4 displays the background soil values for all elements.

3.4.2 SPATIAL AND VERTICAL DISTRIBUTION OF HEAVY METALS IN SOIL

The spatial and vertical distribution of heavy metals showed that metal contamination was not uniform across all boreholes and depths. Figure 3.16 depicts the distance-wise variation in alkalinity, chloride, calcium, and total hardness. The distance-wise fluctuation of Mg, Mn, and K is illustrated in Figure 3.17. Among the analysed metals, Zn, Pb, Mn and Cd showed the most important variation. Zinc reached its maximum concentration of 102 mg/kg at BH-6 at 12 m, lead reached 94 mg/kg at BH-5 at 9 m, manganese reached 510 mg/kg at BH-6 at 12 m, and cadmium reached 9 mg/kg at BH-5 at 9 m.

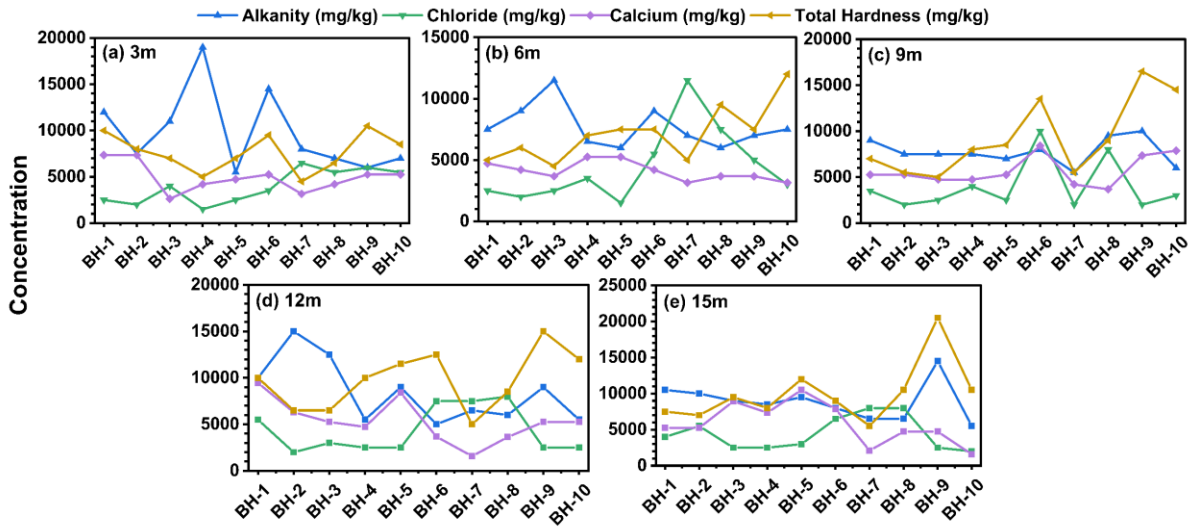


Figure 3.16 Distance-wise assessment of Alkalinity, Cl, Ca, TH assessment in soil samples

These higher values indicate that BH-5 and BH-6 are important metal-enrichment zones in the study area. Figures 3.18 and 3.19 illustrate the distance-wise fluctuation of Zn, Cd, Cu, As, Ag, Pb, Cr, Fr, and Al.

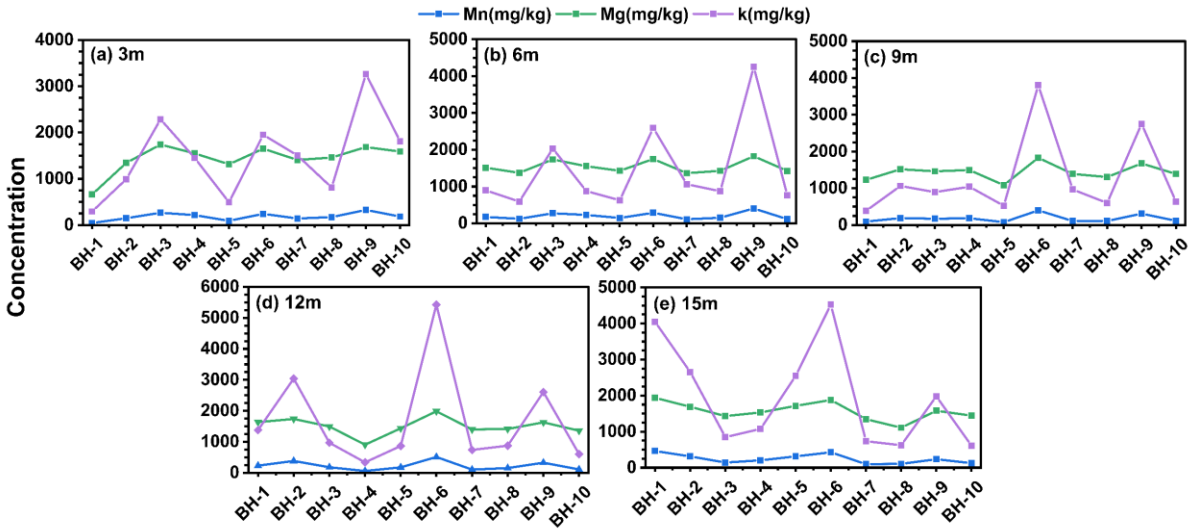


Figure 3.17 Distance-wise assessment of Mn, Mg, K in soil samples

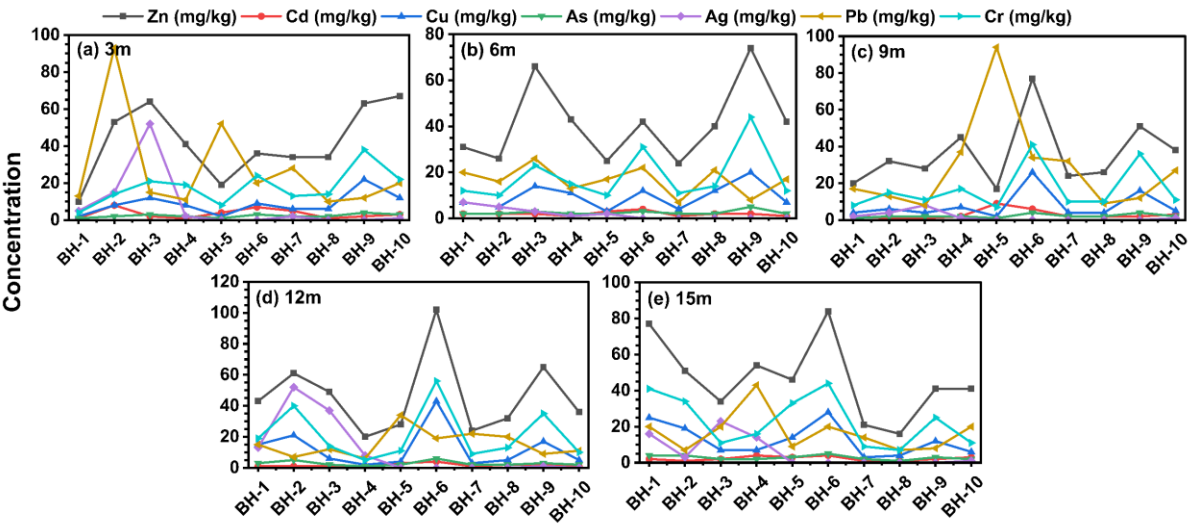


Figure 3.18 Distance-wise assessment of Zn, Cd, Cu, As, Ag, Pb and Cr in soil samples

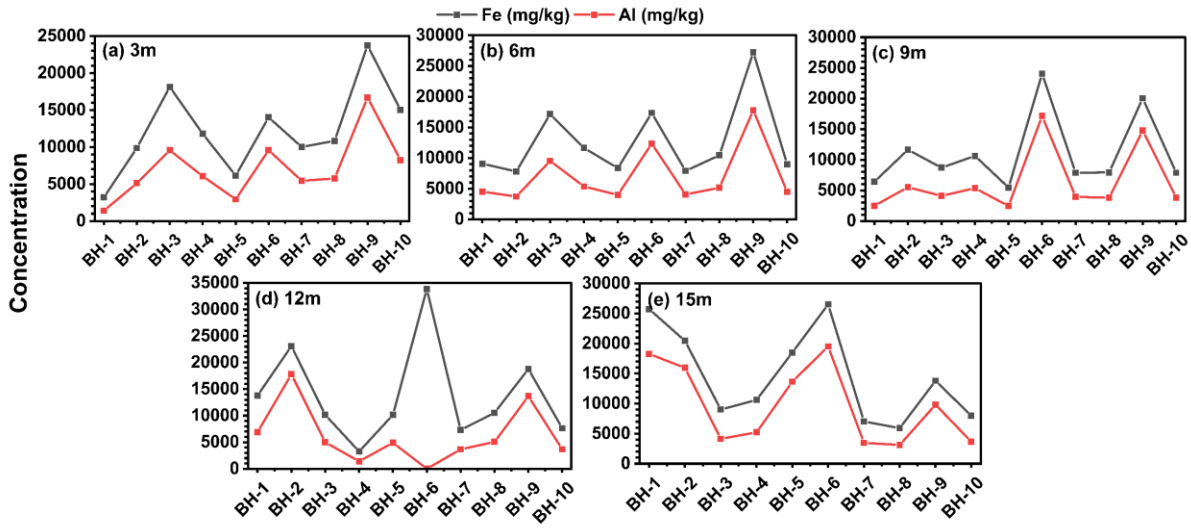


Figure 3.19 Distance-wise assessment of Fe and Al in soil samples

Vertically, higher concentrations were observed not only in shallow soil but also at intermediate depths, especially at 9 m and 12 m. This pattern is important because Bhalswa landfill is an unlined/non-engineered landfill, allowing leachate to percolate downward through the soil profile. The occurrence of maximum Pb and Cd at BH-5 at 9 m, and maximum Zn and Mn at BH-6 at 12 m, indicates that heavy metals have migrated below the surface and accumulated within subsurface soil layers.

Overall, the spatial and vertical distribution confirms that the soil around Bhalswa landfill has been affected by leachate-driven heavy metal migration. The enrichment of Zn, Pb, Mn and Cd at specific boreholes and depths suggests that leachate from the landfill has moved both laterally and vertically, leading to localized contamination within the surrounding soil profile.

3.4.3 STATISTICAL ANALYSIS OF HEAVY METAL CONTAMINATION

Descriptive Statistics of Heavy Metal Concentrations

The descriptive statistics (Table 3.4) reveal substantial variability in metal concentrations within the soil samples collected from the vicinity of the Bhalswa Landfill. Among all analyzed elements, Fe (12,868.8 mg/kg) and Al (7,278.42 mg/kg) exhibited the highest mean concentrations, reflecting their abundance in the soil matrix.

Table 3.5 Statistical Analysis of Heavy Metals Concentration in soil samples

Metal	Mean	SD	Min	Max	Indian Natural Background soil limit (Kuhad et al., 1989 & Gowd et al., 2010)	Indian Limits for soil (Awashthi 2000)	WHO/FAO	CV%
Zn	42.34	19.891	10	102	22.1	300–600	50	**46.98**
Cu	10.12	8.319	2	43	56.5	135–270	36	**82.21**
Cd	2.48	1.843	1	9	Not reported	3–6	0.8	**74.32**
As	2.52	1.147	1	6	Not reported	Not reported	30	**45.52**
Pb	20.94	17.897	6	94	13.1	250–500	85	**85.47**
Cr	19.36	12.647	4	56	114	Not reported	100	**65.33**
Mn	205.3	111.832	45	510	209	Not reported	437	**54.47**
Ag	5.58	11.848	0	52	Not reported	Not reported	Not reported	**212.3**
Fe	12869	6935.05	3195	33854	32015	Not reported	50000	**53.89**
Al	7278.4	5245.23	0	19510	Not reported	Not reported	Not reported	**72.07**
Mg	1496.6	245.421	662	1986	Not reported	Not reported	Not reported	**16.4**
K	1571.5	1237.73	293	5426	Not reported	Not reported	Not reported	**78.76**

All concentration values in mg/kg. CV% = Coefficient of Variation. Colour coding: green < 50%, orange 50–100%, red > 100%. WHO/FAO (World Health Organization/Food and Agriculture Organization) Joint report on Food Hygiene.

The coefficient of variation (CV%) indicates marked differences in the spatial distribution of the elements. Magnesium showed the lowest variability (CV = 16.40%), suggesting a relatively uniform distribution and a dominant geogenic origin. Similarly, Zn (46.98%) and As (45.52%) displayed comparatively low variability, indicating relatively stable concentrations across the study area. In contrast, Cd (74.32%), Cu (82.21%), Pb (85.47%), Cr (65.33%), Al (72.07%), Fe (53.89%), Mn (54.47%), and K (78.76%) exhibited moderate to high variability, reflecting heterogeneous distribution patterns that may be associated with differences in leachate migration,

soil properties, and contaminant accumulation at various depths and borehole locations. The most significant finding is the exceptionally high coefficient of variation of Ag (212.33%), which indicates extreme spatial heterogeneity and the presence of localized enrichment zones or contamination hotspots. Furthermore, the large concentration ranges observed for Fe (3,195–33,854 mg/kg), Al (0–19,510 mg/kg), K (293–5,426 mg/kg), and Pb (6–94 mg/kg) suggest considerable variation among sampling locations. Overall, the elevated variability of Ag, Pb, Cu, and Cd highlights their potential association with landfill leachate influence, whereas the relatively uniform distribution of Mg suggests limited impact from contamination processes.

Spearman Correlation Analysis

Spearman's rank correlation coefficient (Equation 3.2) was used to evaluate monotonic relationships among heavy metals, sampling depth, and distance from the landfill. The analysis was performed on 50 soil samples, and the correlation coefficient values ranged from −1 to +1, where positive values indicate that two variables increase together, negative values indicate an inverse relationship, and values close to zero indicate weak or no monotonic relationship. The resulting correlation coefficients were visualized using a heatmap, and significance levels were shown using star notation: * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

$$\rho = 1 - \frac{6 \sum d^2_i}{n(n^2-1)} \tag{3.2}$$

Where,

ρ = Spearman correlation coefficient

di = difference between the ranks of two variables for each sample

n= number of samples, here n = 50

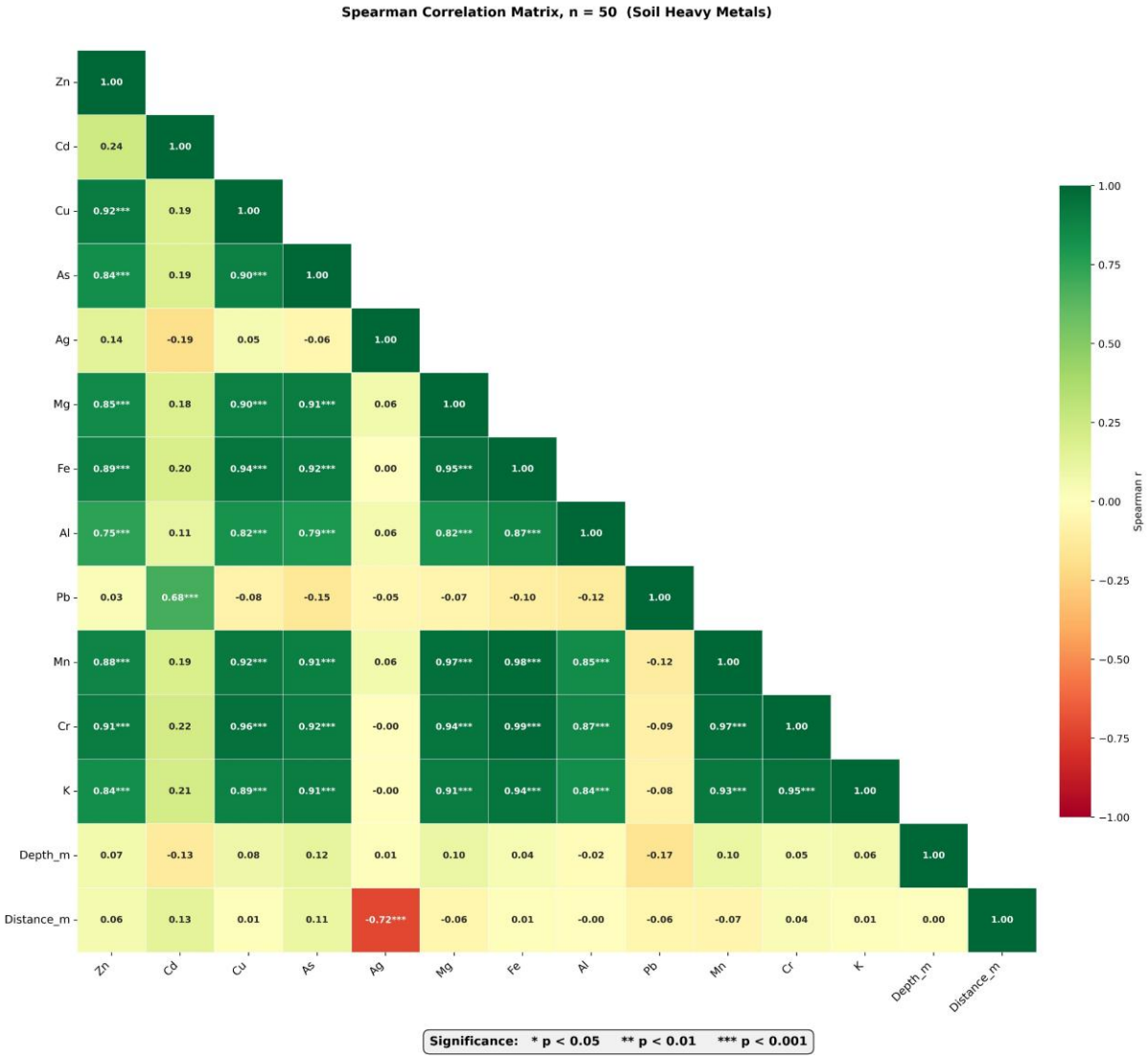


Figure 3.20 Spearman correlation heatmap showing relationships among heavy metals, sampling depth, and distance from landfill in 50 soil samples.

The Spearman correlation matrix (Figure 3.20) showed a strong group of positively correlated elements, including Zn, Cu, As, Mg, Fe, Al, Mn, Cr, and K, with correlation coefficients generally ranging from $\rho = 0.75$ to 0.99 and most relationships significant at $p < 0.001$. The strongest correlations were observed between Fe and Cr ($\rho = 0.99$), Fe and Mn ($\rho = 0.98$), Mn and Cr ($\rho = 0.97$), Cu and Cr ($\rho = 0.96$), and Fe and Mg ($\rho = 0.95$), indicating strong co-occurrence and similar distribution behaviour among these metals. This pattern, together with the contamination observed in leachate, Igeo, and enrichment factor results, suggests that landfill-derived leachate has likely contributed to the enrichment and co-distribution of several metals in the surrounding soils.

However, the strong association among Fe, Al, Mn, Cr, Mg, and K may also reflect natural soil matrix and geochemical control, indicating that both geogenic and landfill-related inputs are influencing the metal distribution.

Sampling depth showed weak correlations with most metals, suggesting that heavy metal variation was not controlled by depth alone across the 50 soil samples. Distance from the landfill also showed weak correlations with most metals; however, Ag displayed a strong and statistically significant negative correlation with distance ($\rho = -0.72$, $p < 0.001$), represented by the prominent red block in the heatmap. This indicates that Ag concentration decreases with increasing distance from the landfill, supporting its role as a sensitive indicator of localized landfill-related contamination.

3.4.4 GEO-ACCUMULATION INDEX AND ENRICHMENT FACTOR ASSESSMENT

Geo-accumulation Index Assessment

The Geo-accumulation Index (Igeo), proposed by Müller (1969), was used to assess the degree of heavy metal contamination in the soil samples using the Equation (3.3). Indian natural background soil values for most metals were taken from Kuhad et al. (1989) and Gowd et al. (2010). The background value for Cd was taken from Indian soil limits reported by Awasthi (2000), while the reference value for As was adopted from WHO/FAO (2007) report. The Igeo classification system and descriptive summary of Igeo values are presented in Tables 3.5 and 3.6, respectively.

$$I_{geo} = \log_2\left(\frac{C_n}{1.5 \times B_n}\right) \quad (3.3)$$

Where:

- C_n: Measured concentration of metal in the soil (mg/kg)

- B_n: Background values of the metal (mg/kg)
- Factor 1.5: Correction factor

The Igeo results indicated that Zn, Cd, and Pb were the major metals contributing to soil contamination in the study area. Zn showed the widest contamination spread, with 32 contaminated samples across different boreholes, reaching Class 2, moderately contaminated, particularly at BH-6 at 12 m, 15 m, and 9 m depths, and at BH-1 at 15 m depth. Cd recorded the highest number of contaminated samples, with 33 samples showing contamination, of which five samples reached Class 3, moderately to heavily contaminated. The highest Cd contamination was observed at BH-5 at 9 m depth with Igeo = 2.907, followed by BH-2 at 3 m depth with Igeo = 2.737. Pb also showed notable contamination, with 22 contaminated samples and Class 3 contamination at BH-5 at 9 m depth with Igeo = 2.258 and BH-2 at 3 m depth with Igeo = 2.243. Mn showed mild contamination in 9 samples, all falling under Class 1, uncontaminated to moderately contaminated, mainly concentrated at BH-6 and BH-9.

Table 3.6 Igeo Classification (Müller, 1969)

Igeo Range	Class	Contamination Level	Description
Igeo ≤ 0	Class 0	Uncontaminated	Natural background levels; no contamination
0 < Igeo ≤ 1	Class 1	Uncontaminated to Moderately Contaminated	Slight enrichment above background
1 < Igeo ≤ 2	Class 2	Moderately Contaminated	Notable enrichment; anthropogenic influence present
2 < Igeo ≤ 3	Class 3	Moderately to Heavily Contaminated	Significant contamination present
3 < Igeo ≤ 4	Class 4	Heavily Contaminated	Strong anthropogenic enrichment
4 < Igeo ≤ 5	Class 5	Heavily to Extremely Contaminated	Very high contamination levels
Igeo > 5	Class 6	Extremely Contaminated	Extreme enrichment — severe pollution

Table 3.7 Descriptive Summary for Geo-Accumulation Index (Igeo)

Metal	B_n (mg/kg)	Mean Igeo	Min Igeo	Max Igeo	Max Igeo Class	Max Sample Contamination Level	All Contaminated Samples (Class ≥ 1)
Zn#	22.1	0.194	-1.729	1.621	**Class 2**	Moderately Contaminated	BH-1: 12m, 15m; BH-2: 3m, 12m, 15m BH-3: 3m, 6m, 12m, 15m BH-4: 3m, 6m, 9m, 15m; BH-5: 15m; BH-6: 3m, 6m, 9m, 12m, 15m; BH-7: 3m BH-8: 3m, 6m; BH-9: 3m, 6m, 9m, 12m, 15m; BH-10: 3m, 6m, 9m, 12m, 15m
Cd@	0.8	0.745	-0.263	2.907	**Class 3**	Moderately to Heavily Contaminated	BH-1: 6m, 9m, 15m BH-2: 3m, 6m; BH-3: 3m, 6m, 15m BH-4: 9m, 15m BH-5: 3m, 6m, 9m, 12m, 15m BH-6: 3m, 6m, 9m, 12m, 15m BH-7: 3m, 9m; BH-8: 6m, 9m, 12m BH-9: 3m, 6m, 9m, 12m, 15m BH-10: 3m, 9m, 15m
Cu#	56.5	-3.482	-5.405	-0.979	**Class 0**	Uncontaminated	None (all Class 0)
As*	30	-4.296	-5.492	-2.907	**Class 0**	Uncontaminated	None (all Class 0)
Pb#	13.1	-0.23	-1.711	2.258	**Class 3**	Moderately to Heavily Contaminated	BH-1: 6m, 15m; BH-2: 3m BH-3: 6m, 15m; BH-4: 9m, 15m BH-5: 3m, 9m, 12m BH-6: 3m, 6m, 9m, 15m BH-7: 3m, 9m, 12m; BH-8: 6m, 12m BH-10: 3m, 9m, 15m
Cr#	114	-3.421	-5.418	-1.61	**Class 0**	Uncontaminated	None (all Class 0)
Fe#	32015	-2.093	-3.91	-0.504	**Class 0**	Uncontaminated	None (all Class 0)
Mn#	209	-0.817	-2.8	0.702	**Class 1**	Uncontaminated to Moderately Contaminated	BH-1: 15m BH-2: 12m BH-5: 15m BH-6: 9m, 12m, 15m BH-9: 3m, 6m, 12m

#Background Values used from Kuhad et al., 1989 & Gowd et al., 2010; @Awasthi (2000); *WHO/FAO, 2007 reference limit.

In contrast, Cu, As, Cr, and Fe showed no contamination across the analysed samples, as their Igeo values remained at or below 0, corresponding to Class 0 or uncontaminated status. Overall, the Igeo assessment suggests that Cd, Zn, and Pb are the primary metals of concern, with BH-5 at 9 m and BH-2 at 3 m emerging as the most critically contaminated locations. This pattern indicates localized contamination, likely influenced by landfill leachate migration, subsurface transport pathways, and site-specific soil conditions.

Enrichment Factor Assessment

The Enrichment Factor (EF), proposed by Buat-Ménard and Chesselet (1979), was used as given in Equation 3.4, to evaluate the degree of anthropogenic enrichment of heavy metals in the soil samples. EF was calculated using Fe as the reference element, as Fe is commonly used to normalize metal concentrations against natural soil background variation. Background values for Zn, Cu, Pb, Cr, Mn, and Fe were taken from Kuhad et al. (1989) and Gowd et al. (2010), while Cd and As values were taken from WHO/FAO (2007) reference limits. The detailed summary of EF and EF classification proposed by Sutherland (2000), is included in Tables 3.7 and 3.8.

$$EF = \frac{(C_x/C_{ref})_{\text{sample}}}{(B_x/B_{ref})_{\text{background}}} \quad (3.4)$$

Where:

- C_x: Concentration of the examined element in the sample
- C_{ref}: Concentration of a reference element (Here, we assumed Fe) in the sample
- B_x: Concentration of examine element in the background sample
- B_{ref}: Concentration of the reference element in the background sample

The EF results showed clear enrichment of Zn, Cd, Pb, and Mn in the soil samples. Cd recorded the highest mean EF value of 9.663, with values ranging from 1.733 to 66.013, indicating

significant enrichment and, in some samples, extremely high enrichment. A total of 48 out of 50 Cd samples exceeded the EF > 2 threshold, showing enrichment in 96% of the samples. Pb showed a mean EF of 5.477, with a maximum EF of 42.105, indicating significant enrichment and localized extremely high enrichment; 39 out of 50 samples (78%) were enriched with Pb. Zn also showed significant enrichment, with a mean EF of 5.010 and all 50 samples exceeding the EF > 2 threshold, indicating widespread Zn enrichment across the study area. Mn showed moderate enrichment, with a mean EF of 2.436 and 48 out of 50 samples exceeding the enrichment threshold. In contrast, Cu, As, and Cr showed deficiency to minimal enrichment, with mean EF values of 0.401, 0.221, and 0.402, respectively, and none of their samples exceeded the EF > 2 threshold. Overall, the EF assessment indicates that Zn, Cd, Pb, and Mn are the main enriched metals in the study area, suggesting substantial anthropogenic contribution, most likely associated with landfill leachate influence, waste decomposition, and localized subsurface migration, whereas Cu, As, and Cr appear to remain close to natural/background levels.

Table 3.8 Descriptive Summary for Enrichment Factor

Enrichment Factor (EF) — Summary Table							
Metal	Background (mg/kg)	Mean EF	Min EF	Max EF	Std EF	Enrichment Level	@Enriched Samples (EF > 2)
Zn#	22.1	5.01	3.506	8.904	1.238	Significant enrichment	50/50 (100%)
Cd*	0.8	9.663	1.733	66.013	10.229	Significant enrichment	48/50 (96%)
Cu#	56.5	0.401	0.184	0.72	0.123	Deficiency to minimal	0/50 (0%)
As*	30	0.221	0.166	0.334	0.043	Deficiency to minimal	0/50 (0%)
Pb#	13.1	5.477	0.717	42.105	6.846	Significant enrichment	39/50 (78%)
Cr#	114	0.402	0.304	0.523	0.058	Deficiency to minimal	0/50 (0%)
Mn#	209	2.436	1.879	3.014	0.275	Moderate enrichment	48/50 (96%)

#Background Values used from Kuhad et al., 1989 & Gowd et al., 2010.
*WHO/FAO, 2007 reference limit.
@EF > 2 mean, the concentration of that metal in the sample is more than 2 times higher than the natural background level, which indicates that anthropogenic (human/industrial) sources are contributing to its presence in the soil.

The coefficient of variation (CV%) indicates marked differences in the spatial distribution of the elements. Magnesium showed the lowest variability (CV = 16.40%), suggesting a relatively uniform distribution and a dominant geogenic origin. Similarly, Zn (46.98%) and As (45.52%) displayed comparatively low variability, indicating relatively stable concentrations across the study area. In contrast, Cd (74.32%), Cu (82.21%), Pb (85.47%), Cr (65.33%), Al (72.07%), Fe (53.89%), Mn (54.47%), and K (78.76%) exhibited moderate to high variability, reflecting heterogeneous distribution patterns that may be associated with differences in leachate migration, soil properties, and contaminant accumulation at various depths and borehole locations. The most significant finding is the exceptionally high coefficient of variation of Ag (212.33%), which indicates extreme spatial heterogeneity and the presence of localized enrichment zones or contamination hotspots. Furthermore, the large concentration ranges observed for Fe (3,195–33,854 mg/kg), Al (0–19,510 mg/kg), K (293–5,426 mg/kg), and Pb (6–94 mg/kg) suggest considerable variation among sampling locations. Overall, the elevated variability of Ag, Pb, Cu, and Cd highlights their potential association with landfill leachate influence, whereas the relatively uniform distribution of Mg suggests limited impact from contamination processes.

Table 3.9 Enrichment Factor Classification

EF Value	Class	Enrichment Level
EF < 2	**Class 1**	Deficiency to minimal enrichment
2 – 5	**Class 2**	Moderate enrichment
5 – 20	**Class 3**	Significant enrichment
20 – 40	**Class 4**	Very high enrichment
EF > 40	**Class 5**	Extremely high enrichment

The EF class distribution across all boreholes shows that Cu, As, and Cr consistently fall under Class 1, indicating deficiency to minimal enrichment and suggesting that these metals remain close to natural/background levels. Zn mainly falls under Class 2 and Class 3, representing moderate to significant enrichment across several boreholes, while Mn is mostly classified as Class 2, indicating moderate enrichment in most samples. Cd and Pb show the highest enrichment severity, with several samples falling under Class 3, indicating significant enrichment, and selected samples reaching Class 4 and Class 5, representing very high and extremely high enrichment, respectively. The occurrence of higher EF classes for Cd and Pb, especially in specific boreholes, indicates localized anthropogenic enrichment, likely influenced by landfill-related contamination and subsurface migration pathways. Overall, the EF class pattern suggests that Cd, Pb, Zn, and Mn are the main enriched metals, whereas Cu, As, and Cr show minimal enrichment across the study area.

3.5 GEOTECHNICAL BEHAVIOUR OF LEACHATE-IMPACTED SOIL

The geotechnical behaviour of leachate-impacted soil was evaluated using 50 samples collected from ten boreholes at depths of 3 m, 6 m, 9 m, 12 m and 15 m. The analysis included index properties, compaction characteristics, permeability and shear strength parameters. The results indicate that the soil around the Bhalswa landfill is mainly sandy in nature, but localized variation in fines, moisture content, compaction response and shear strength was observed, especially in boreholes closer to the landfill and at deeper depths.

3.5.1 INDEX AND COMPACTION PROPERTIES

The natural water content varied from 6.10% to 21.29%. The lowest value was recorded at BH-10 at 3 m, while the highest value was observed at BH-1 at 15 m. In most boreholes, natural water content increased with depth; for example, it increased from 11.36% at 3 m to 21.29% at 15 m in BH-1, and from 6.10% at 3 m to 13.80% at 15 m in BH-10. This increasing trend indicates higher moisture retention in deeper soil layers.

Based on the IS 1498:1970 soil classification, the soil profile around the Bhalswa landfill was predominantly sandy, but distinct depth-wise and distance-wise variations were observed. The samples were mainly classified as Poorly Graded Sand (SP), Poorly Graded Sand with Silt (SP-SM), Silty Sand with Gravel (SM), Clayey Sand with Gravel (SC) and Low-Plasticity Clay/Sandy Clay with Gravel (CL). The dominance of SP and SP-SM indicates that the study area is largely composed of sandy alluvial deposits; however, the occurrence of fine-rich and plastic soil layers at certain boreholes suggests localized alteration in the subsurface soil profile. Figure 3.21(a, b, c, d, e, f, g, h, i, j) depicts the particle size distribution curve for all 50 samples collected from BH 1 to BH 10.

In the boreholes located closer to the landfill, clear vertical variation was observed. At BH-1, located at 306.81 m, the soil changed from Silty Sand with Gravel (SM) at 3 m to Poorly Graded Sand with Silt (SP-SM) at 6 m, Poorly Graded Sand (SP) at 9 m, again SP-SM at 12 m, and finally Low-Plasticity Clay/Sandy Clay with Gravel (CL) at 15 m.

The change from predominantly sandy soil in the upper and intermediate layers to a fine-rich plastic layer at 15 m is important, as this layer contained 70% silt + clay and showed plasticity with LL = 30.60%, PL = 22.50% and PI = 8.10%. Similarly, BH-2, located at 619.27 m, was classified as SP from 3 m to 12 m, but changed to Clayey Sand with Gravel (SC) at 15 m, where the fines content and plasticity increased. These changes at deeper depths in BH-1 and BH-2

indicate that the soil near the landfill is not uniform with depth and that leachate percolation may have contributed to the modification of soil layers, fines accumulation and plastic behaviour in deeper layers.

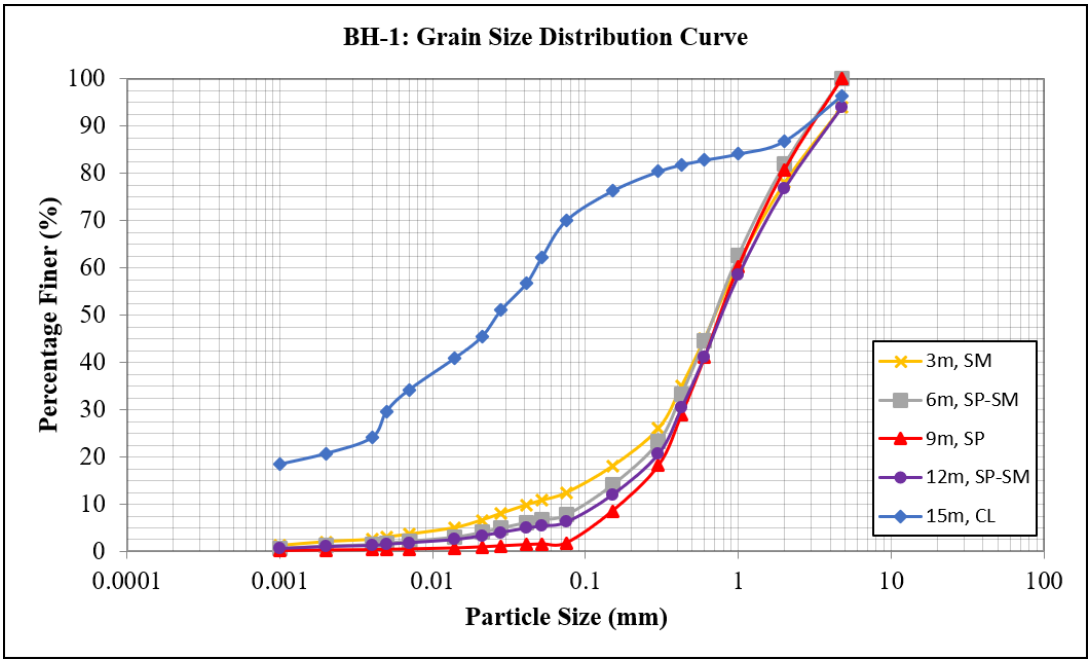


Figure 3.21 (a) Particle Size distribution for BH-1

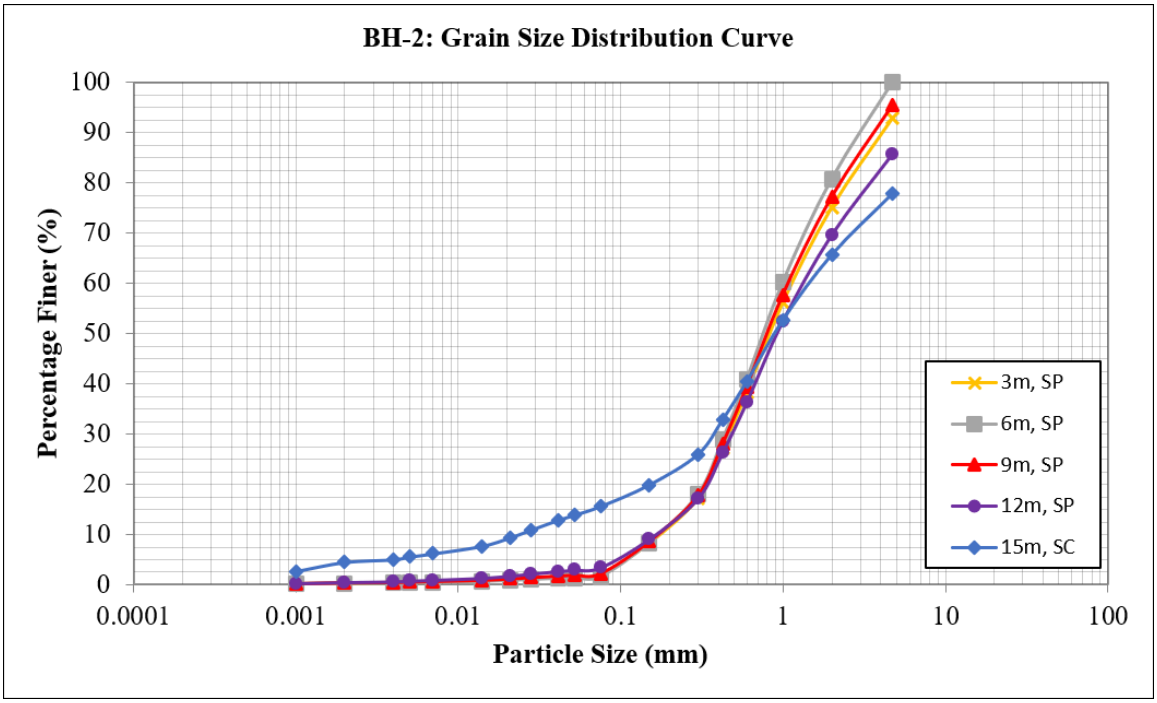


Figure 3.21 (b) Particle Size distribution for BH-2

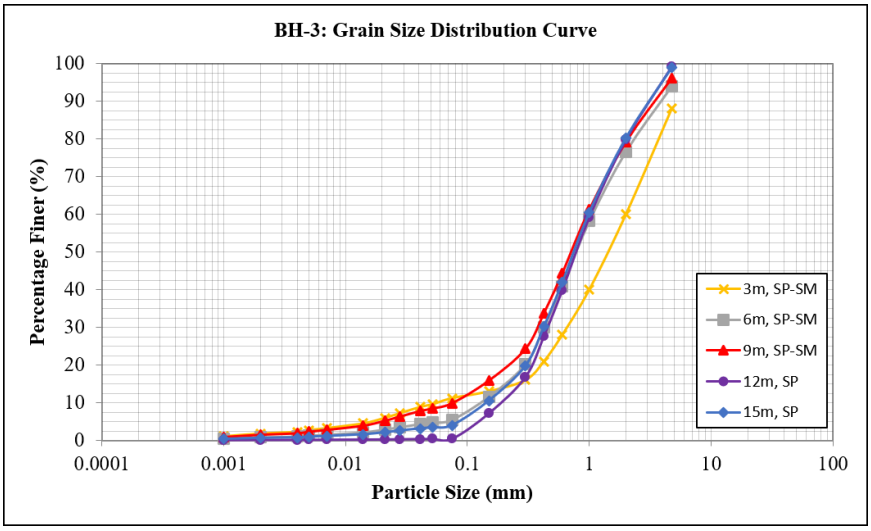


Figure 3.21 (c) Particle Size distribution for BH-3

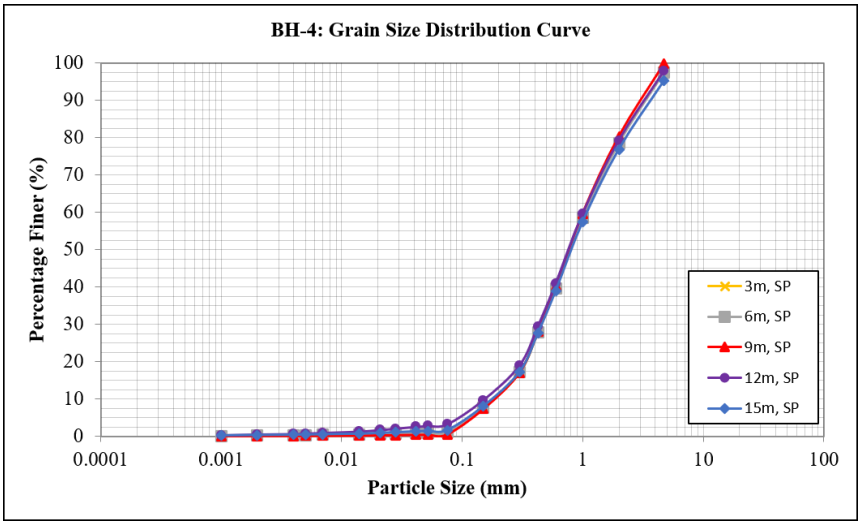


Figure 3.21 (d) Particle Size distribution for BH-4

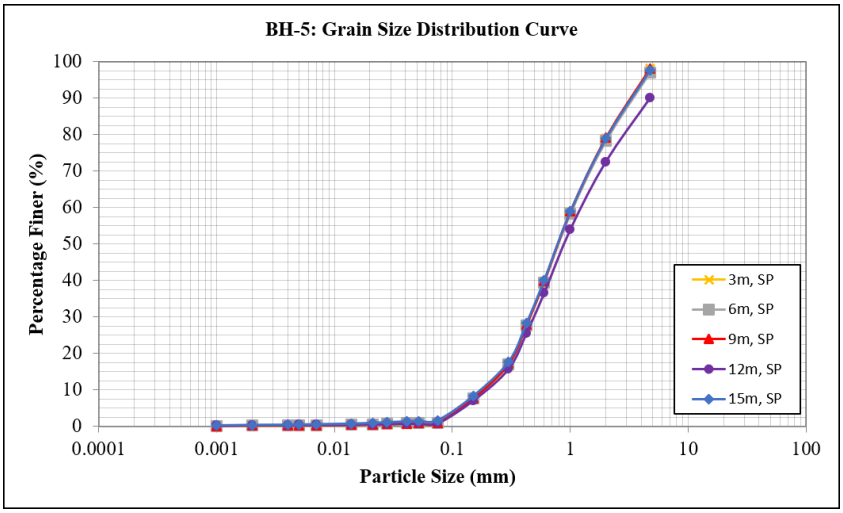


Figure 3.21 (e) Particle Size distribution for BH-5

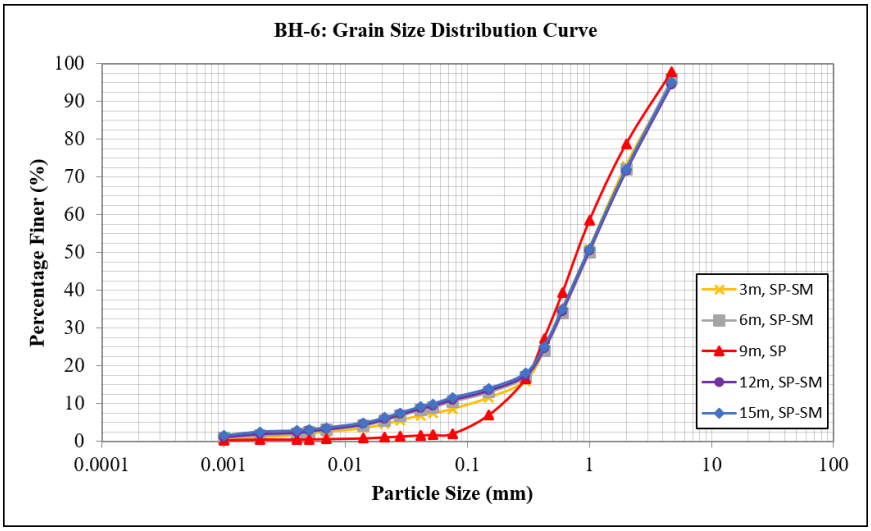


Figure 3.21 (f) Particle Size distribution for BH-6

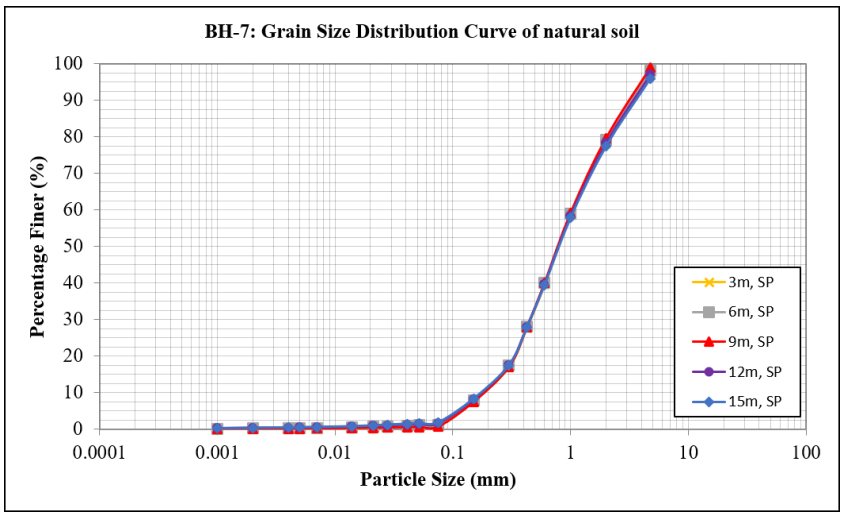


Figure 3.21 (g) Particle Size distribution for BH-7

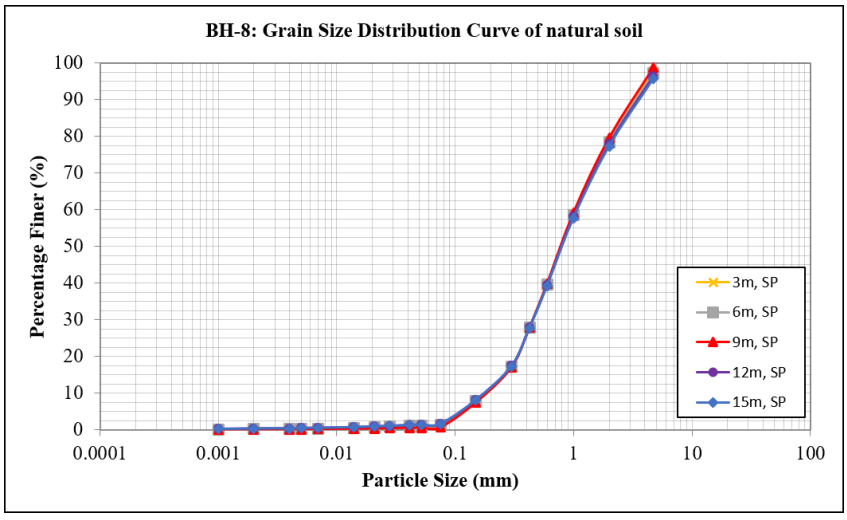


Figure 3.21 (h) Particle Size distribution for BH-8

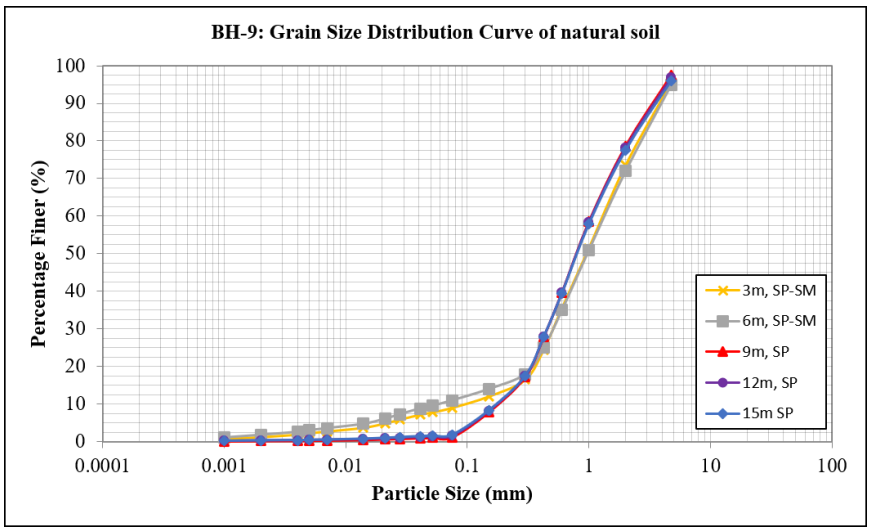


Figure 3.21 (i) Particle Size distribution for BH-9

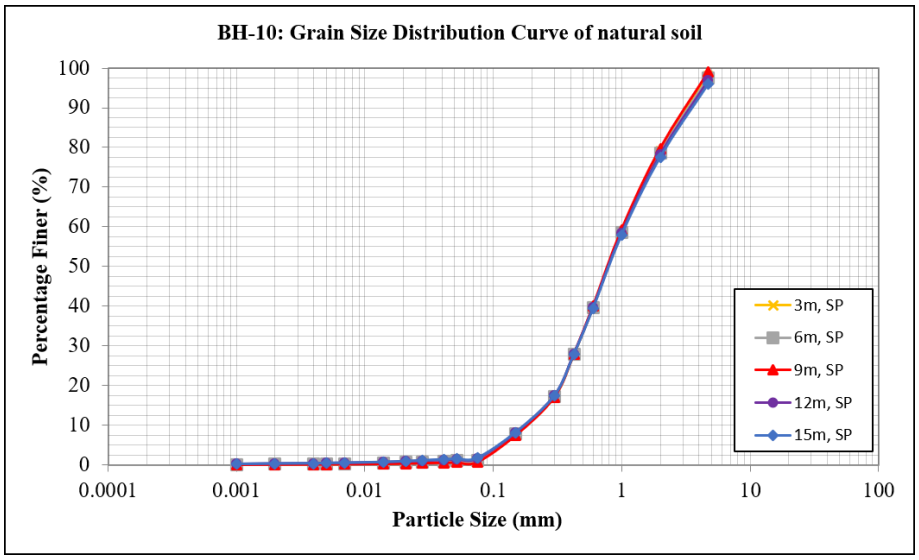


Figure 3.21 (j) Particle Size distribution for BH-10

In the intermediate-distance boreholes, the soil classification was comparatively more stable but still showed localized variation. BH-3, located at 728.58 m, showed SP-SM from 3 m to 9 m, followed by SP at 12 m and 15 m, indicating a reduction in silty fraction with depth. BH-4 at 829.51 m and BH-5 at 882.48 m were classified as SP throughout all depths from 3 m to 15 m, suggesting that these boreholes retained predominantly clean sandy behaviour. However, BH-6, located at 1239.69 m, again showed alternating behaviour, with SP-SM at 3 m, 6 m, 12 m and 15

m, and SP only at 9 m. This indicates that even at a greater distance, localized silty layers are present, which may influence moisture retention, compaction response and permeability.

The farther boreholes generally showed more uniform sandy behaviour. BH-7 at 1361.47 m, BH-8 at 1810.56 m and BH-10 at 2012.93 m were classified as SP at all depths, indicating relatively stable poorly graded sandy strata with very low fines content and non-plastic behaviour. BH-9 at 1889.03 m showed SP-SM at 3 m and 6 m, followed by SP from 9 m to 15 m, indicating minor silty influence in the shallow layers but cleaner sandy behaviour at greater depths. Among all boreholes, BH-10 showed the most consistent SP classification throughout the profile and was therefore considered the least affected borehole for comparison of compaction changes.

Overall, the classification pattern shows that the soil profile is dominantly sandy, but the boreholes nearer to the landfill, particularly BH-1 and BH-2, show clear depth-wise changes towards fine-rich and plastic soil at 15 m. This behaviour is significant because Bhalswa landfill is an unlined/non-engineered landfill, allowing leachate to percolate downward through the subsurface. The occurrence of CL and SC layers at deeper depths near the landfill, along with SP-SM layers in some boreholes, suggests that leachate interaction and fines accumulation may have modified the soil behaviour locally. In contrast, the farther boreholes, especially BH-10, show uniform SP classification, indicating comparatively lower geotechnical alteration with distance from the landfill.

Most of the soil samples were non-plastic. Plasticity was observed mainly at deeper contaminated horizons, particularly BH-1 at 15 m, with LL = 30.60%, PL = 22.50% and PI = 8.10%, and BH-2 at 15 m, with LL = 20.65%, PL = 14.00% and PI = 6.65%. The occurrence of plasticity at these depths indicates the presence of finer soil fractions and possible alteration of soil behaviour due to leachate interaction.

Specific gravity varied narrowly from 2.60 to 2.75, indicating limited variation in mineral density. The lowest value was observed at BH-1 at 9 m, while the highest value was recorded at BH-3 at 12 m.

The compaction behaviour was evaluated using Optimum Moisture Content (OMC) and Maximum Dry Density (MDD). The percentage change in OMC and MDD for each layer was calculated with respect to the corresponding depth of BH-10, which was considered the least affected borehole due to its maximum distance from the landfill and its uniform Poorly Graded Sand (SP) classification throughout the profile. Therefore, the values at BH-10 were treated as the reference condition for assessing leachate-related changes in compaction behaviour.

The OMC values of the soil samples ranged from 9.01% to 18.00%, while MDD varied from 16.20 to 19.88 kN/m³. The lowest OMC was recorded at BH-10 at 9 m, whereas the highest OMC was recorded at BH-2 at 15 m. The lowest MDD was observed at BH-1 at 6 m, while the highest MDD was recorded at BH-10 at 15 m. In comparison with BH-10, most boreholes showed positive percentage change in OMC and negative percentage change in MDD, indicating that the leachate-impacted soils generally required higher moisture for compaction but achieved lower dry density. Figure 3.22 depicts the OMC and MDD plot for BH-1 to BH-10 soil samples.

Depth-wise variation was clearly visible in the boreholes closer to the landfill. At BH-1, OMC increased from 9.28% at 3 m to 14.50% at 15 m, while MDD reduced from 17.56 kN/m³ at 3 m to 17.30 kN/m³ at 15 m. The MDD reduction was highest at BH-1 at 6 m, with a percentage change of −15.98%, and the deepest layer at BH-1 at 15 m also showed a reduction of −12.98%. This behaviour is consistent with the presence of a fine-rich Low-Plasticity Clay/Sandy Clay with Gravel (CL) layer at 15 m. At BH-2, the most significant change was observed at 15 m, where OMC reached 18.00% with a percentage increase of 56.52%, while MDD reduced to 16.60 kN/m³

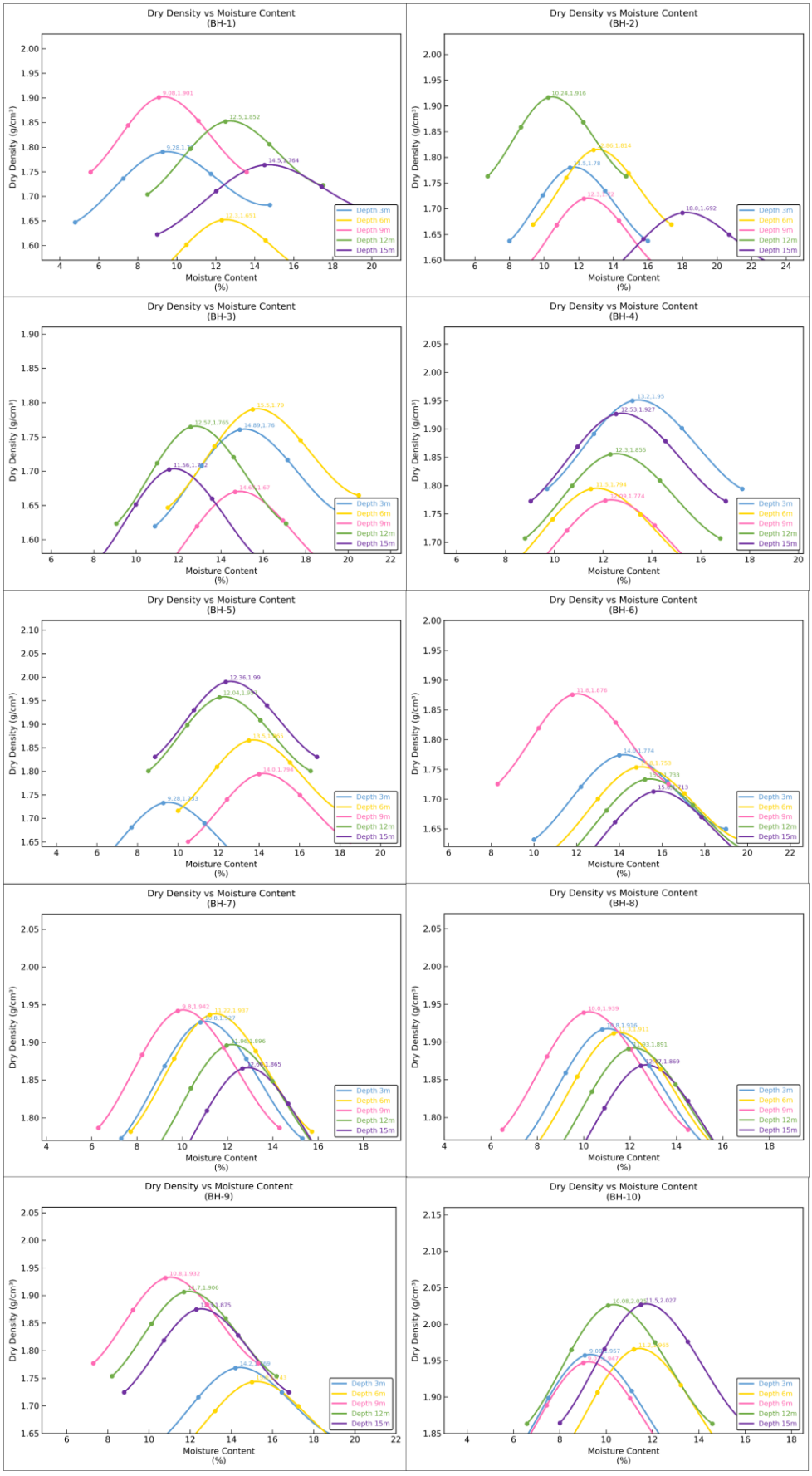


Figure 3.22 Compaction Curves for BH-1 to BH-10

with a percentage change of −16.50%. This corresponds to the Clayey Sand with Gravel (SC) layer and indicates strong modification of compaction behaviour at depth.

At intermediate distances, BH-3 showed high OMC increase at 3 m and 9 m, with percentage changes of 63.99% and 62.82%, respectively. The MDD reduction was also notable, especially at BH-3 at 9 m and 15 m, with percentage changes of −14.24% and −16.00%. BH-4 and BH-5 showed comparatively less severe MDD reduction, particularly at deeper depths. For example, BH-5 at 15 m showed only −1.81% change in MDD, suggesting relatively better compactability compared with the more affected layers. However, BH-5 at 9 m showed a high OMC increase of 55.38%, indicating localized moisture demand despite its sandy classification. A distinct compaction response was also observed at BH-6, located at 1239.69 m. The OMC increased considerably at BH-6 at 3 m and 12 m, with percentage changes of 54.19% and 50.79%, respectively. The MDD reduction was highest at BH-6 at 12 m and 15 m, with percentage changes of −14.44% and −15.49%. This indicates that even at this distance, leachate-affected or fine-influenced layers may show reduced dry density and higher moisture requirement. In contrast, the farther boreholes BH-7, BH-8 and BH-10 generally showed more stable compaction behaviour, with smaller changes in MDD in several layers. BH-10 showed 0% change in both OMC and MDD at all depths because it was used as the reference borehole.

Overall, the compaction results indicate that leachate-impacted and fine-rich layers require higher moisture content to achieve compaction and generally develop lower maximum dry density. The negative percentage change in MDD reflects a reduction in compactability, which may be attributed to increased fines content, moisture retention and leachate-induced alteration of soil particle arrangement. The most affected layers were observed at BH-1 at 6 m and 15 m, BH-2 at 15 m, BH-3 at 9 m and 15 m, and BH-6 at 12 m and 15 m.

3.5.2 PERMEABILITY AND SHEAR STRENGTH CHARACTERISTICS

The permeability and shear strength characteristics of the soil samples showed clear variation with depth and distance from the Bhalswa landfill. The coefficient of permeability ranged from 4.30 × 10^{-6} m/s to 17.80 × 10^{-6} m/s**,** indicating low to moderate permeability behaviour. In general, the sandy samples classified as Poorly Graded Sand (SP**)** showed higher permeability, whereas samples containing higher fines and plasticity showed lower permeability. The lowest permeability was recorded at BH-1 at 15 m**,** where the soil was classified as Low-Plasticity Clay/Sandy Clay with Gravel (CL) and contained the highest silt + clay fraction. Low permeability was also observed at BH-2 at 15 m**,** which was classified as Clayey Sand with Gravel (SC)**.** These results indicate that the presence of fines and plastic soil fractions reduced the drainage capacity of the soil.

Depth-wise variation in permeability was more prominent in boreholes closer to the landfill. At BH-1**,** permeability decreased from 13.20 × 10^{-6} m/s at 3 m to 4.30 × 10^{-6} m/s at 15 m**.** Similarly, at BH-2**,** permeability reduced from 14.10 × 10^{-6} m/s at 3 m to 4.90 × 10^{-6} m/s at 15 m. This reduction at deeper depths corresponds with the increase in fines and change in soil classification from sandy soil to clayey/fine-rich soil. In contrast, the farther boreholes, especially BH-10, maintained comparatively higher and more uniform permeability values throughout the depth profile, consistent with its uniform SP classification. Higher permeability values were observed in sandy layers such as BH-5 at 3 m**,** BH-8 at 9 m**,** BH-10 at 9 m**,** and BH-7 at 9 m, indicating that clean sandy strata allowed easier movement of water or leachate through the soil matrix. Figure 3.23 depicts the direct shear test plots for BH-1 to BH-10.

The shear strength parameters also varied across the boreholes and depths. Cohesion ranged from 1.00 kPa to 18.00 kPa, while the angle of internal friction varied from 27.00° to 33.40°**.** The highest cohesion was recorded at BH-1 at 15 m, followed by BH-2 at 15 m**,** BH-6 at 15 m**,** and BH-6 at

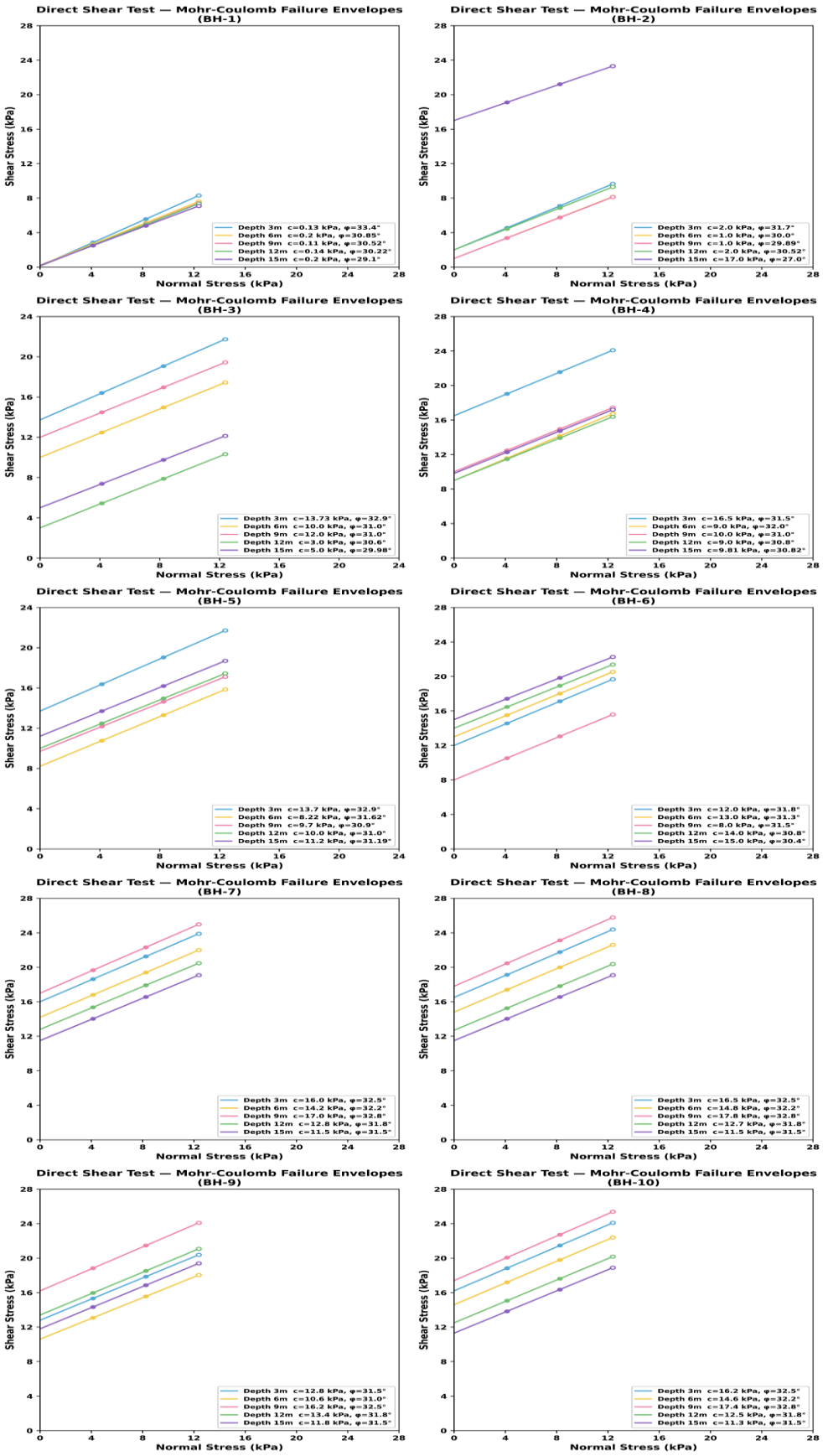


Figure 3.23 Direct Shear test plots for BH-1 to BH-10

12 m**.** These higher cohesion values are associated with increased fines content, greater moisture retention and the presence of clayey or silty fractions. In contrast, lower cohesion values were recorded in clean sandy samples, particularly in the upper and intermediate sandy layers of BH-2**,** where the soil remained mostly non-plastic.

The angle of internal friction was generally higher in sandy samples and lower in fine-rich layers. The highest friction angle of 33.40° was recorded at BH-1 at 3 m, where the soil was classified as Silty Sand with Gravel (SM) with dominant sand and gravel fractions. Most SP-classified samples showed friction angles above 31°, indicating friction-dominated shear resistance. However, the friction angle decreased at depths where fines and plasticity increased. The lowest friction angle of 27.00° was recorded at BH-2 at 15 m**,** while BH-1 at 15 m also showed a comparatively lower value of 29.10°. These layers also showed higher cohesion, confirming that the shear strength behaviour changed from predominantly frictional in sandy layers to more cohesive and fine-controlled in deeper affected layers.

Overall, the permeability and shear strength results show that leachate-impacted soil behaviour varies significantly with soil classification, depth and distance from the landfill. The nearer boreholes, especially BH-1 and BH-2, showed reduced permeability and modified shear strength at deeper depths due to the presence of fine-rich layers. The increase in cohesion and reduction in friction angle at these depths indicate that leachate interaction, fines accumulation and higher moisture retention have altered the engineering behaviour of the soil. In contrast, the farther boreholes, particularly BH-10, retained more uniform sandy behaviour with comparatively higher permeability and friction-dominated shear strength.

3.6 SEM–EDX-BASED MICROSTRUCTURAL AND ELEMENTAL ANALYSIS OF SOIL

SEM–EDX-based microstructural and elemental analysis was performed on ten soil samples collected at 3 m depth from boreholes BH-1 to BH-10 in the vicinity of the Bhalswa landfill to assess the influence of landfill leachate on soil morphology and elemental composition. The samples were analysed using a Thermo Fisher Apreo Scanning Electron Microscope equipped with an EDAX APEX EDX detector, operated at an accelerating voltage of 15.00 kV with the ETD detector in Standard mode. Quantitative elemental analysis was carried out using the eZAF correction algorithm at 99% confidence, ensuring reliable estimation of major and trace elements present in the soil matrix. Each borehole sample was examined at four magnifications, namely 1,000×, 5,000×, 10,000×, and 20,000×, across two analytical sessions, allowing observation of soil morphology from broader particle arrangement to finer surface-level features.

Table 3.10 Cross-Borehole EDX Summary

Element / Range	Al (wt%)	Pb (wt%)	Cd (wt%)	Cr (wt%)	Cu (wt%)	Zn (wt%)	As (wt%)
BH-1	44.5–98.8	0–53.4	0.7–3.1	0–15.2	0–13.4	0–1.0	0
BH-2	24.9–92.7	0–72.4	2.1–7.1	0–3.5	0–12.1	0–3.7	0–0.8
BH-3	35.3–82.8	6–23.1	0.9–8.7	1.3–22.2	0–20.3	0.2–3.2	0.3–1.0
BH-4	53.8–78.4	6.2–25.4	1.3–8.4	0.3–7.0	0–9.0	0–2.8	0.4–0.7
BH-5	49.6–84.6	4.4–37.6	5.0–6.7	0.9–6.0	0.1–5.6	0–1.7	0.1–0.5
BH-6	60.9–73.7	5.5–21.3	3.6–7.3	1.0–3.4	0–3.3	0–1.3	0.2–0.8
BH-7	57.2–82.0	2.4–33.9	2.8–7.3	0–4.7	0–12.2	0–3.3	0–1.7
BH-8	37.6–82.1	5.5–53.5	1.0–6.4	1.4–7.8	0–12.2	0–2.3	0–1.1
BH-9	34.3–83.1	3.6–46.4	1.5–7.0	0.9–4.1	1.4–11.7	0–2.3	0.1–1.6
BH-10	72.2–83.1	3.6–17.1	5.9–7.0	0.9–1.7	2.8–3.7	0.6–0.9	0.1–0.4

SEM imaging was used to study particle texture, aggregation, pore spaces, surface roughness, and possible contaminant deposition, while EDX analysis helped identify elemental signatures such as oxygen, silicon, aluminium, iron, calcium, magnesium, sodium, potassium, and other leachate-associated elements. Overall, the SEM–EDX investigation provides microstructural and elemental

evidence for evaluating spatial variation in soil contamination and understanding the impact of Bhalswa landfill leachate on surrounding soils. Table 3.9 summarises the measured elemental ranges for the seven principal elements detected across all ten boreholes. All values are in weight percent (wt%). Ranges reflect minimum and maximum values recorded across all analytical spots per borehole.

Bore Hole-1

BH-1 shows a heterogeneous microstructural and elemental character under SEM–EDX analysis (Figure 3.24 – 3.25). SEM images indicate a dominant Al-rich mineral matrix with smooth to slightly rough planar surfaces, cleavage-like particle edges, scattered fine fragments, and angular detrital grains distributed over the surface. The EDX observations show that contamination is not uniformly spread across the matrix but occurs as localized, high-density inclusions and discrete particles. Bright Pb–Cd-rich agglomerates are observed on selected grain surfaces, suggesting localized deposition or physical trapping of heavy-metal-bearing particles rather than continuous surface coating. In addition, isolated Cr–Cu-bearing angular phases are present, indicating the occurrence of separate contaminant-bearing fragments within the sample. The absence of a consistent arsenic signal suggests that As was not a major contaminant phase in BH-1 at the analysed depth. Overall, BH-1 reflects a matrix-dominated sample with spatially uneven metal association, where localized Pb–Cd and Cr–Cu hotspots indicate patchy interaction with landfill-derived contaminant particles.

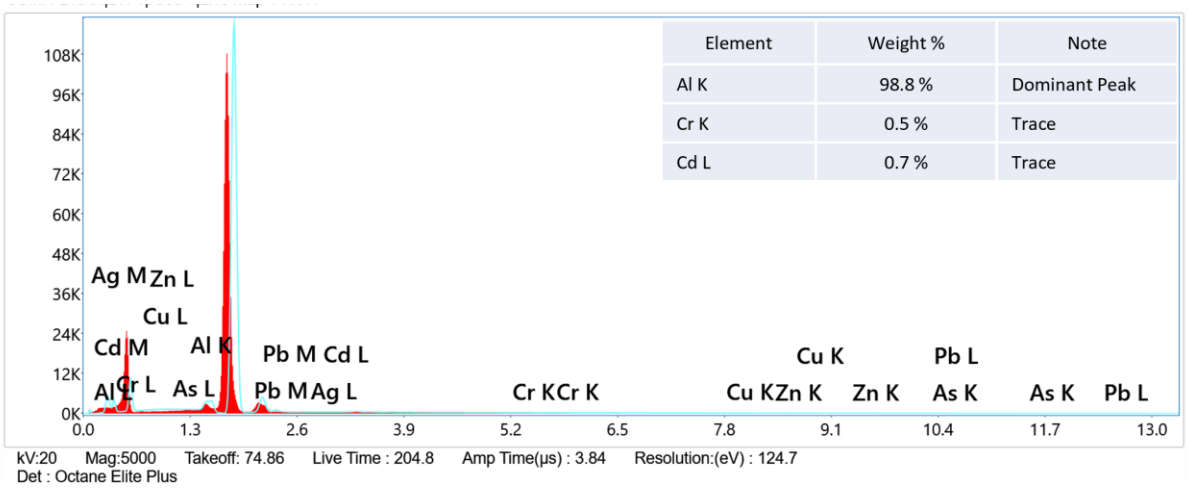


Figure 3.24 EDS Spectrum — BH-1 (Highest Peak Area, EDAX APEX™)

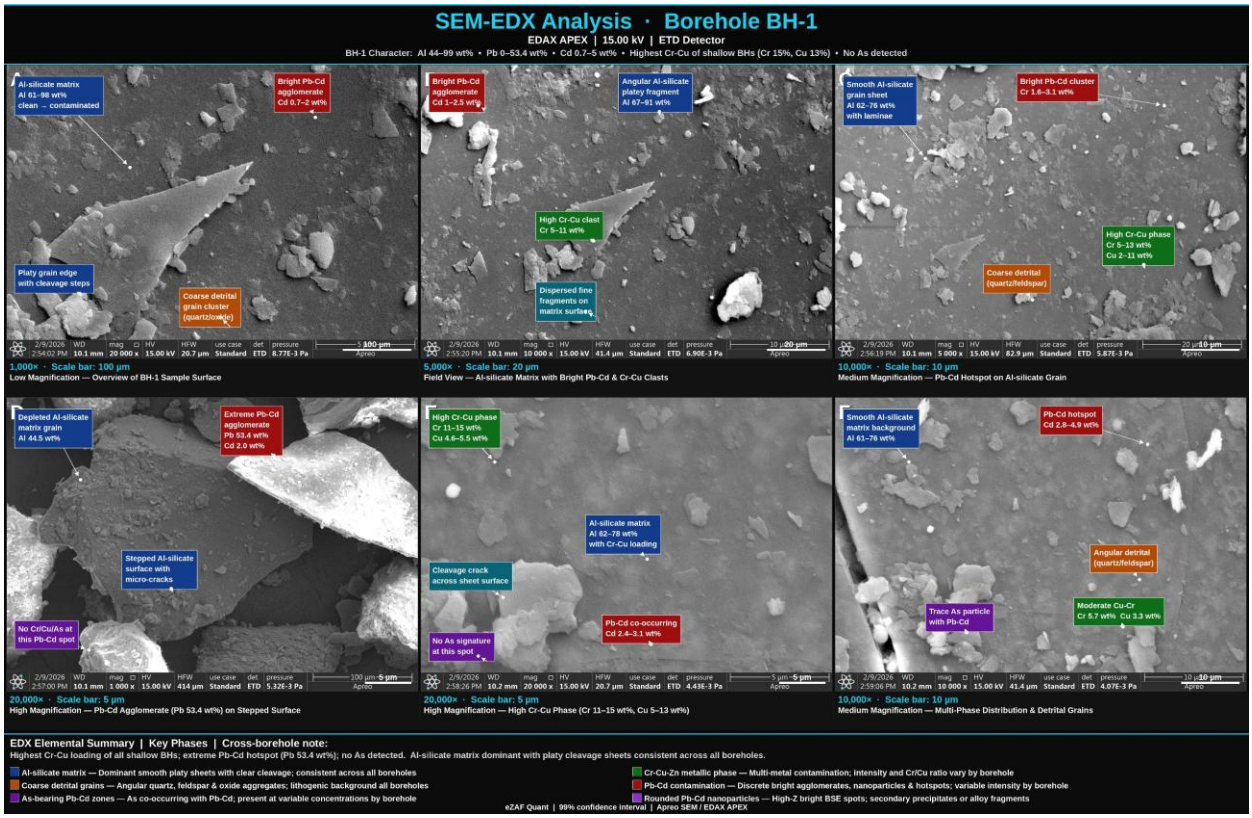


Figure 3.25 SEM observation of BH-1 soil sample

Bore Hole-2

BH-2 (Figure 3.26 – 3.27). shows a highly heterogeneous SEM–EDX character, with an Al-rich mineral matrix interrupted by several localized heavy-metal-bearing particles and agglomerates.

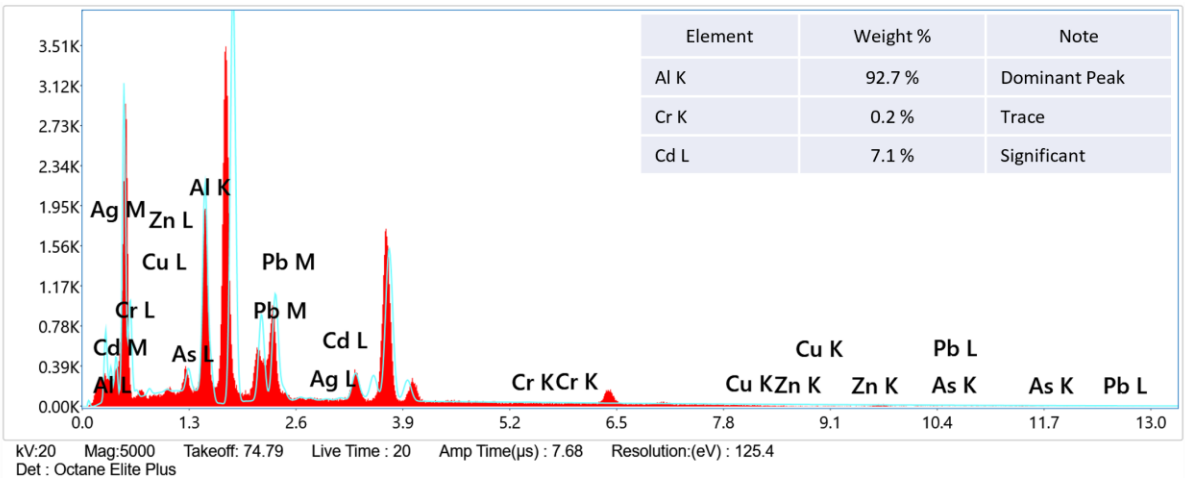


Figure 3.26 EDS Spectrum — BH-2 (Highest Peak Area, EDAX APEX™)

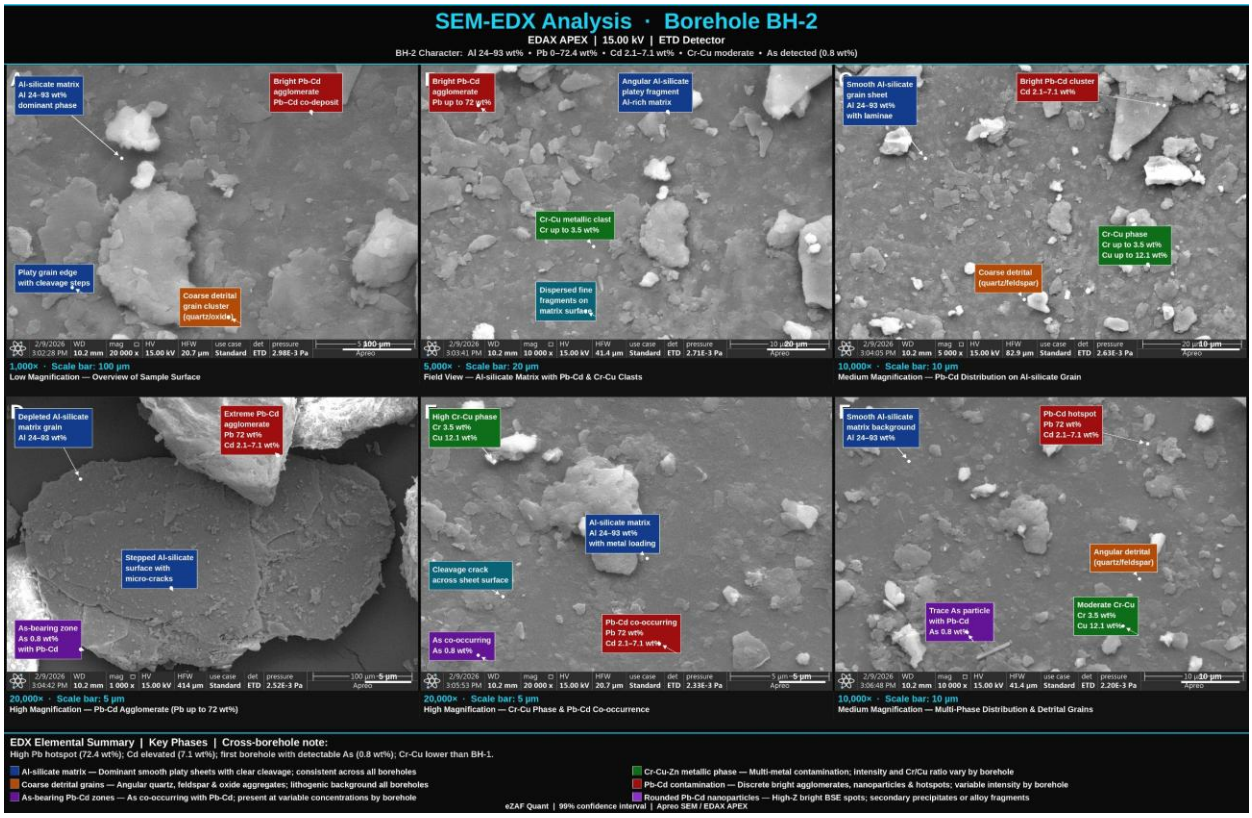


Figure 3.27 SEM observation of BH-2 soil sample

SEM images reveal smooth to stepped matrix surfaces, scattered fine fragments, angular detrital grains, and bright high-density inclusions distributed unevenly across the analysed areas. The most prominent feature of BH-2 is the presence of intense Pb–Cd-rich agglomerates and hotspots, which

occur as discrete bright particles on the matrix surface rather than as uniform coatings, suggesting localized deposition or physical trapping of contaminant phases. In addition to Pb–Cd association, Cu–Cr and Cu–Zn-bearing phases are also observed as angular and dispersed fragments, indicating the presence of multiple contaminant sources or mixed landfill-derived waste particles. The occurrence of arsenic in selected spots, particularly near Pb–Cd-rich areas, suggests limited co-association of As with the metal-bearing contaminant phase. Overall, BH-2 reflects stronger and more complex localized contamination than a uniform matrix alteration, with Pb–Cd-dominant hotspots and secondary Cu–Cr–Zn phases indicating patchy but significant interaction with landfill-derived leachate particles.

Bore Hole-3

BH-3, SEM–EDX (Figure 3.28 – 3.29) analysis indicates a clear shift from Pb-dominant contamination toward Cr–Cu-associated metallic phases. The SEM micrographs show rough, angular, and fragmented particles distributed over the matrix, with several bright high-density inclusions representing localized metal-rich zones. Unlike BH-1 and BH-2, where Pb–Cd particles were more prominent, BH-3 is mainly characterized by Cr–Cu-bearing phases with associated Cd and Zn, suggesting the presence of complex metallic or alloy-type fragments derived from landfill waste. These contaminant particles appear texturally distinct from the surrounding matrix, indicating that they are present as separate embedded fragments rather than as uniform surface coatings. Pb–Cd clusters are still observed, but they are less dominant compared to the Cr–Cu-rich particles. The limited occurrence of arsenic in selected areas further suggests localized association with metal-bearing phases. Overall, BH-3 reflects concentrated metallic contamination, particularly linked to Cr–Cu–Cd–Zn-bearing fragments, indicating localized deposition of landfill-derived industrial or metallic waste components within the soil matrix.

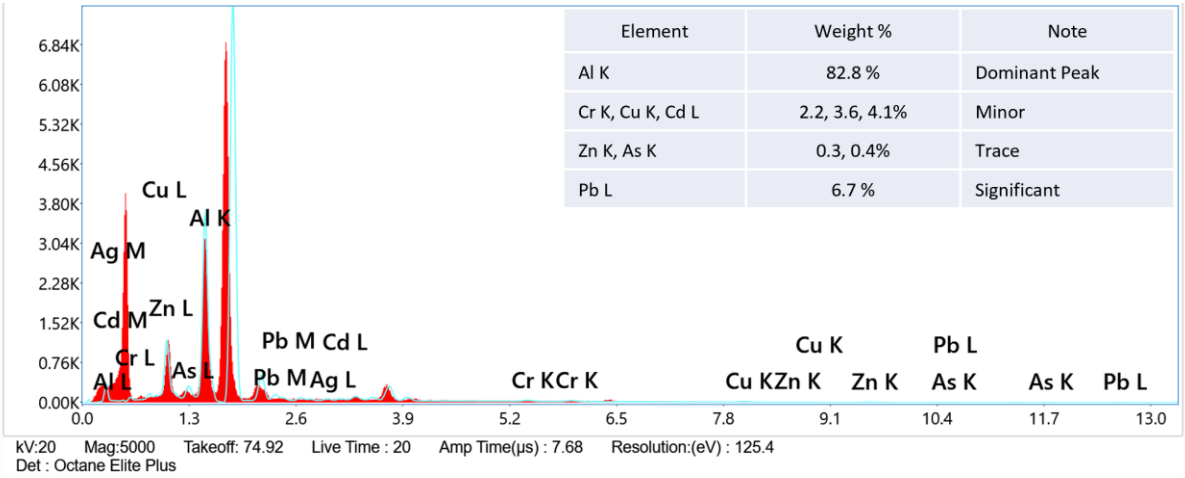


Figure 3.28 EDS Spectrum — BH-3 (Highest Peak Area, EDAX APEX™)

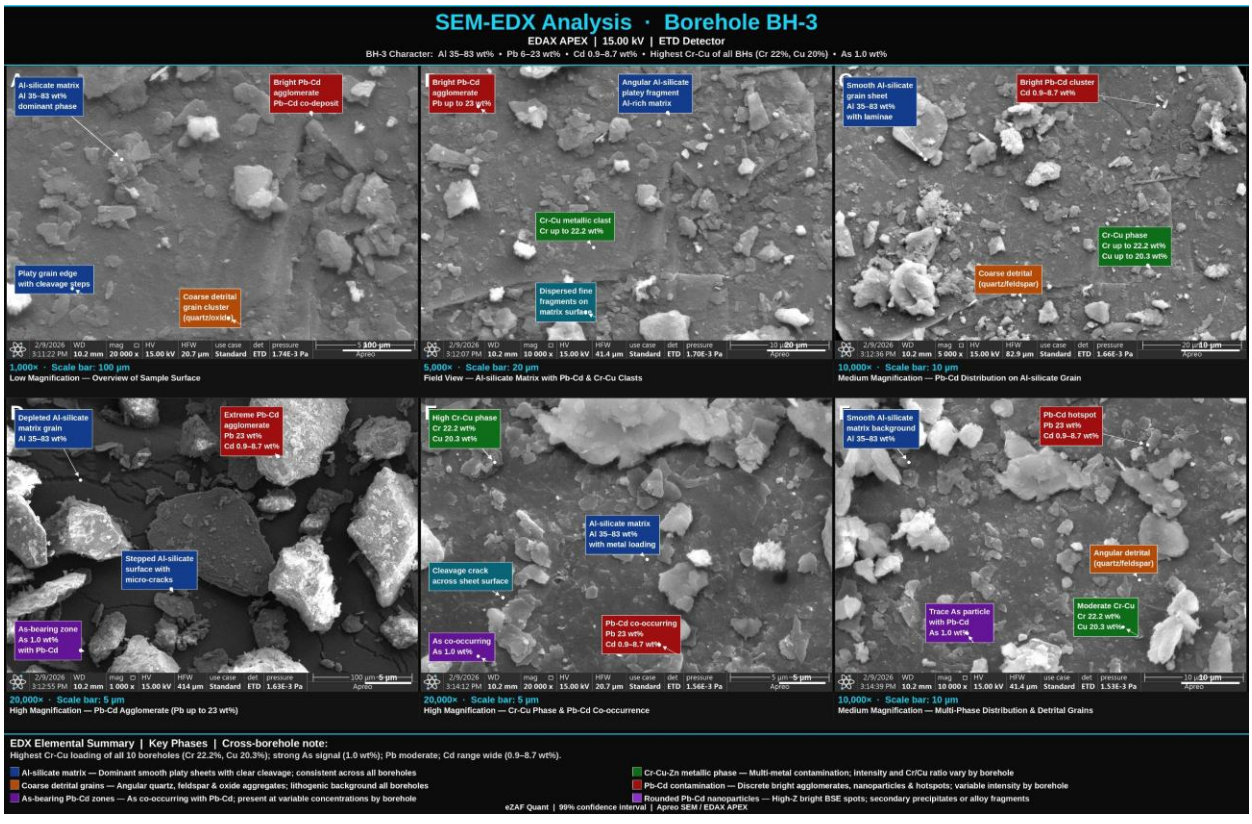
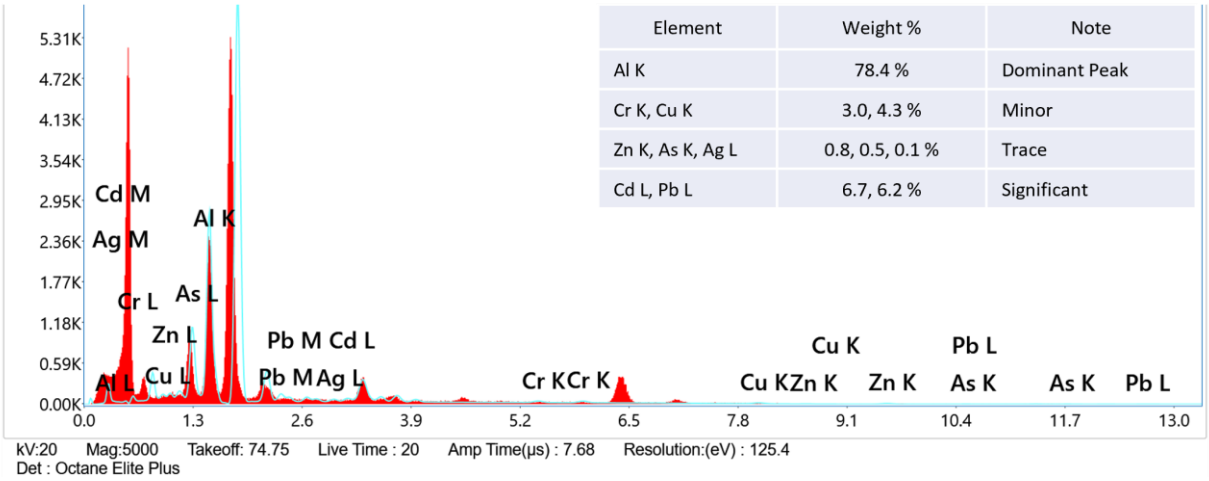


Figure 3.29 SEM observation of BH-3 soil sample

Bore Hole-4

BH-4 (Figure 3.30 – 3.31) indicates a transitional stage in the SEM–EDX profile, where the severe localized metal loading observed in the earlier boreholes begins to reduce, but contaminant-bearing phases remain clearly present.



Element	Weight %	Note
Al K	78.4 %	Dominant Peak
Cr K, Cu K	3.0, 4.3 %	Minor
Zn K, As K, Ag L	0.8, 0.5, 0.1 %	Trace
Cd L, Pb L	6.7, 6.2 %	Significant

Figure 3.30 EDS Spectrum — BH-4 (Highest Peak Area, EDAX APEX™)

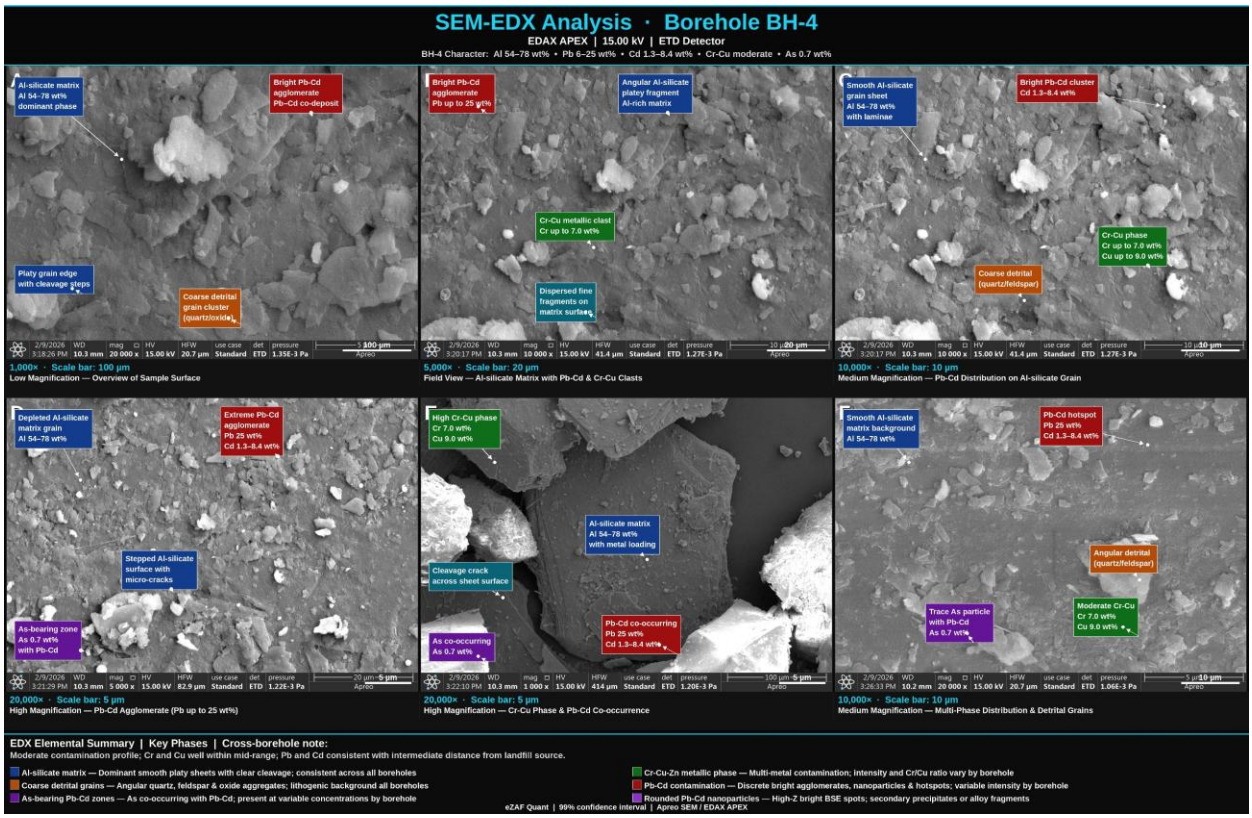


Figure 3.31 SEM observation of BH-4 soil sample

The SEM micrographs show relatively more stable Al-rich matrix surfaces with smooth to stepped particle features, scattered fine fragments, and angular detrital grains. A distinctive feature of BH-4 is the occurrence of elongate, needle-like crystalline particles associated with Cr–Cu-bearing zones, suggesting possible formation of secondary mineral precipitates rather than only direct deposition of primary landfill waste particles. Pb–Cd-rich agglomerates are still present, but they appear more moderate and dispersed compared to the intense hotspots observed in BH-1 to BH-3. The consistent occurrence of arsenic across the analysed spots indicates that As has become more widely associated with the contaminant-bearing matrix at this borehole. Overall, BH-4 reflects a shift from highly localized primary metal fragments toward more transformed and dispersed contamination, with evidence of secondary Cr–Cu-rich crystalline phases formed through interaction with landfill-derived leachate.

Bore Hole-5

BH-5 shows evidence of a declining but still active contamination influence in the SEM–EDX (Figure 3.32 – 3.33) observations. The SEM micrographs display broad Al-rich matrix surfaces with smoother particle areas, stepped grain edges, scattered fine debris, and angular detrital fragments, indicating that the original matrix remains more visually continuous compared to the highly disrupted zones observed in earlier boreholes. However, localized Pb–Cd-rich clusters are still clearly present as bright, discrete particles on the grain surfaces, showing that heavy-metal-bearing material continues to occur in isolated zones rather than as a uniform coating. A notable feature of BH-5 is the more consistent presence of Cd across the analysed areas, suggesting that cadmium may be comparatively more mobile and more evenly distributed at this location, while Pb remains concentrated in selected hotspots. Cr–Cu-bearing phases are also observed but appear less intense and more dispersed, indicating attenuation of these contaminants with distance or

progressive immobilization before reaching this borehole. Trace arsenic is present in selected areas, reflecting limited association with the contaminant-bearing phases.

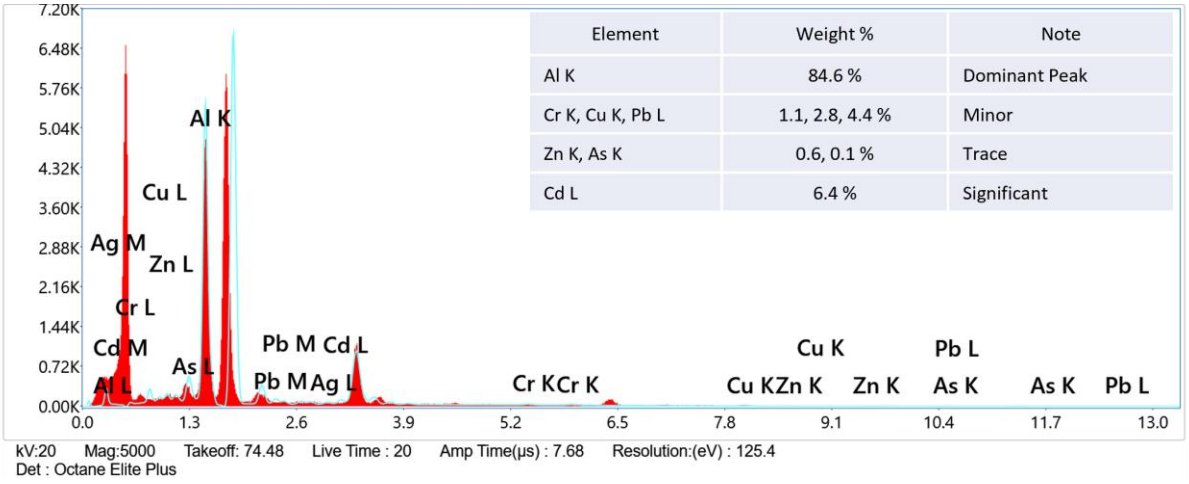


Figure 3.32 EDS Spectrum — BH-5 (Highest Peak Area, EDAX APEX™)

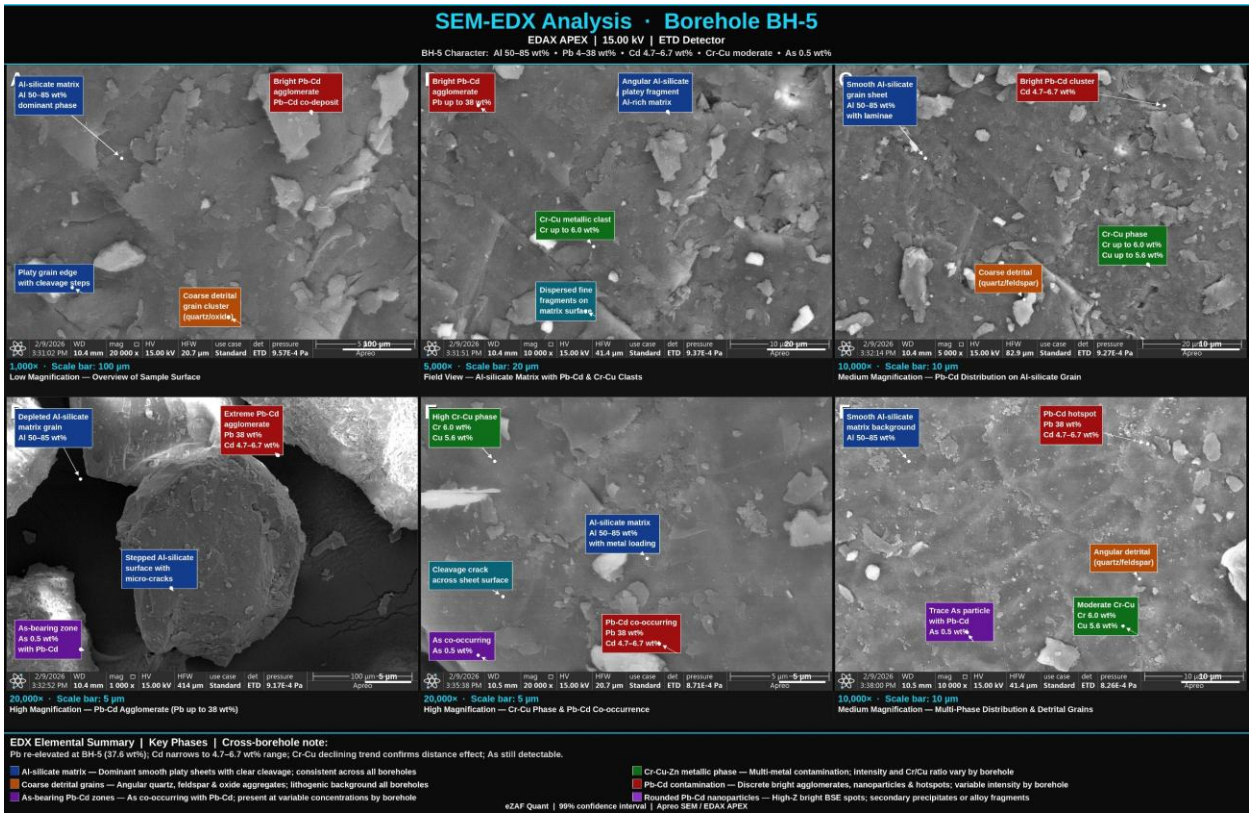


Figure 3.33 SEM observation of BH-5 soil sample

Overall, BH-5 represents an intermediate stage where the matrix appears comparatively more preserved, but residual Pb–Cd contamination and weaker Cr–Cu signatures confirm continued interaction with landfill-derived leachate and metal-bearing particles.

Bore Hole-6

BH-6 presents a more evenly distributed contamination pattern compared to the earlier boreholes, with SEM–EDX (Figure 3.34 – 3.35) observations showing a relatively uniform Al-rich matrix affected by moderate metal association rather than extreme localized hotspots.

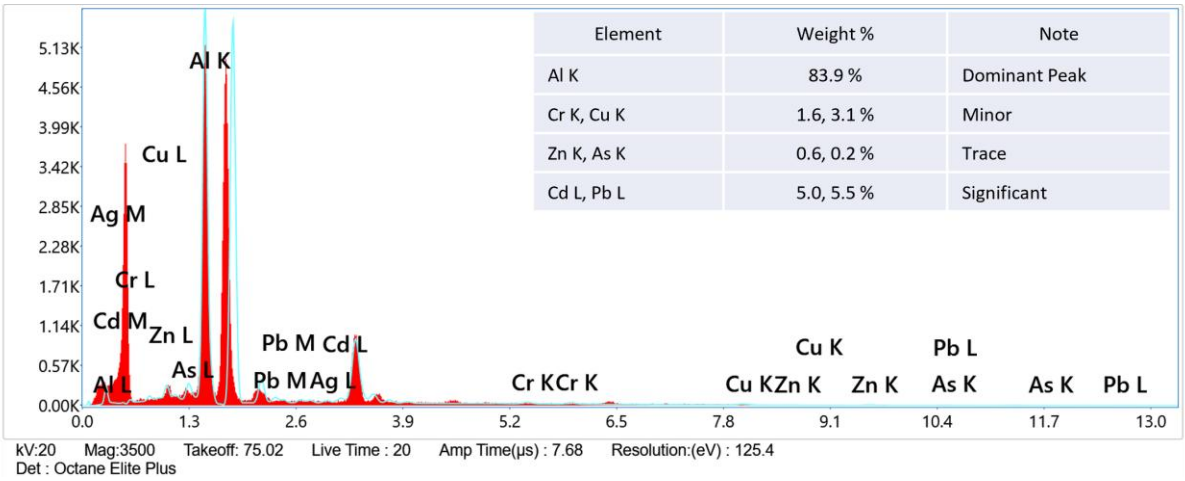


Figure 3.34 EDS Spectrum — BH-6 (Highest Peak Area, EDAX APEX™)

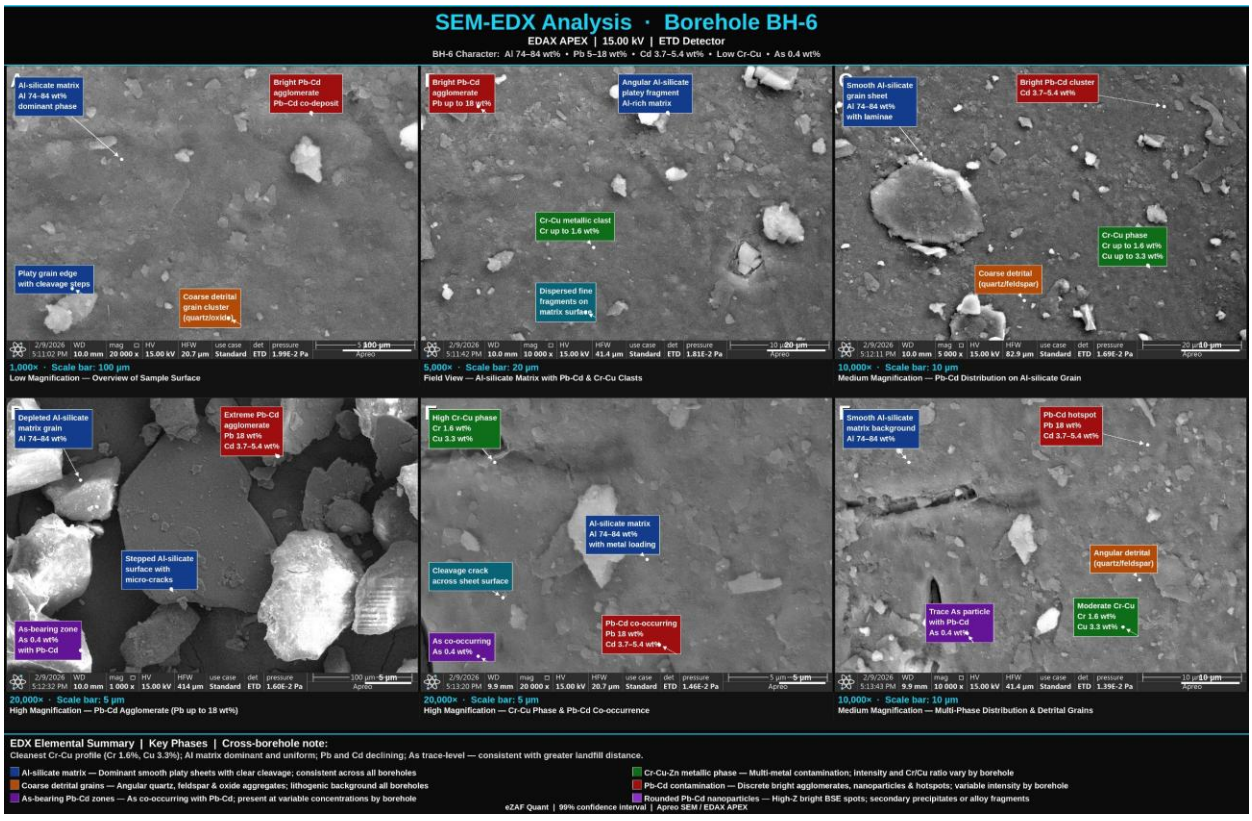


Figure 3.35 SEM observation of BH-6 soil sample

The SEM micrographs display smooth to slightly stepped matrix surfaces, scattered fine particles, angular detrital grains, and limited bright inclusions distributed across the analysed areas. Pb–Cd-bearing particles are still present, but they appear as moderate clusters and isolated surface deposits rather than intense, dominant agglomerates. A key feature of BH-6 is the consistent occurrence of arsenic across the analysed spots, suggesting that As has become more regularly associated with Pb–Cd-bearing phases at this location, possibly due to co-precipitation or stabilization within secondary contaminant phases. In contrast, Cr–Cu signatures are weaker than those observed in the earlier boreholes, indicating attenuation of these metals during transport through the soil matrix. Overall, BH-6 reflects a moderately contaminated zone where leachate-derived metals are more evenly dispersed, with persistent Pb–Cd–As association and reduced Cr–Cu influence.

Bore Hole-7

BH-7 (Figure 3.36 – 3.37) shows a renewed increase in metal association after the relatively moderate pattern observed in BH-5 and BH-6, indicating that contamination at this location is not simply declining with distance.

The SEM micrographs show a rough and uneven matrix surface with fractured, stepped grains, fine particulate deposits, and several bright high-density inclusions attached to or embedded within the particle surfaces. Pb–Cd-rich clusters are again clearly visible as localized agglomerates, while Cu–Cr–Zn-bearing phases occur as separate angular fragments within the matrix, suggesting the presence of mixed metallic waste-derived particles. The co-occurrence of arsenic with these metal-rich phases, along with the notable Ag signature seen in selected spots, indicates a more complex contaminant assemblage than in the preceding boreholes. This pattern suggests that BH-7 may be influenced by a secondary contaminant input, preferential leachate pathway, or localized accumulation of dense metal-bearing particles. Overall, BH-7 represents an anomalous zone of

renewed contamination, marked by Pb–Cd clusters, Cu–Cr–Zn phases, arsenic association, and scattered Ag-bearing signatures within a visibly roughened microstructure.

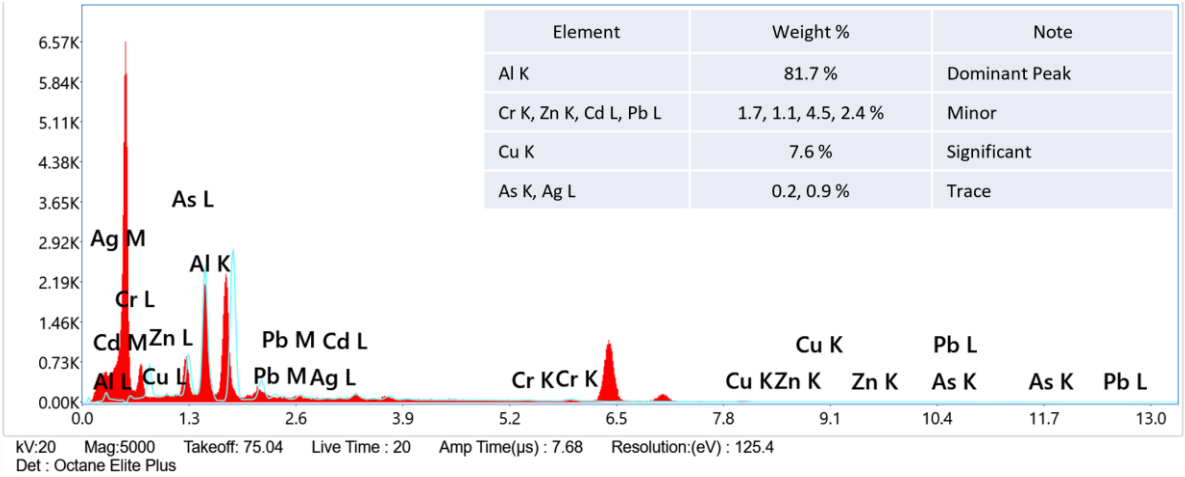


Figure 3.36 EDS Spectrum — BH-7 (Highest Peak Area, EDAX APEX™)

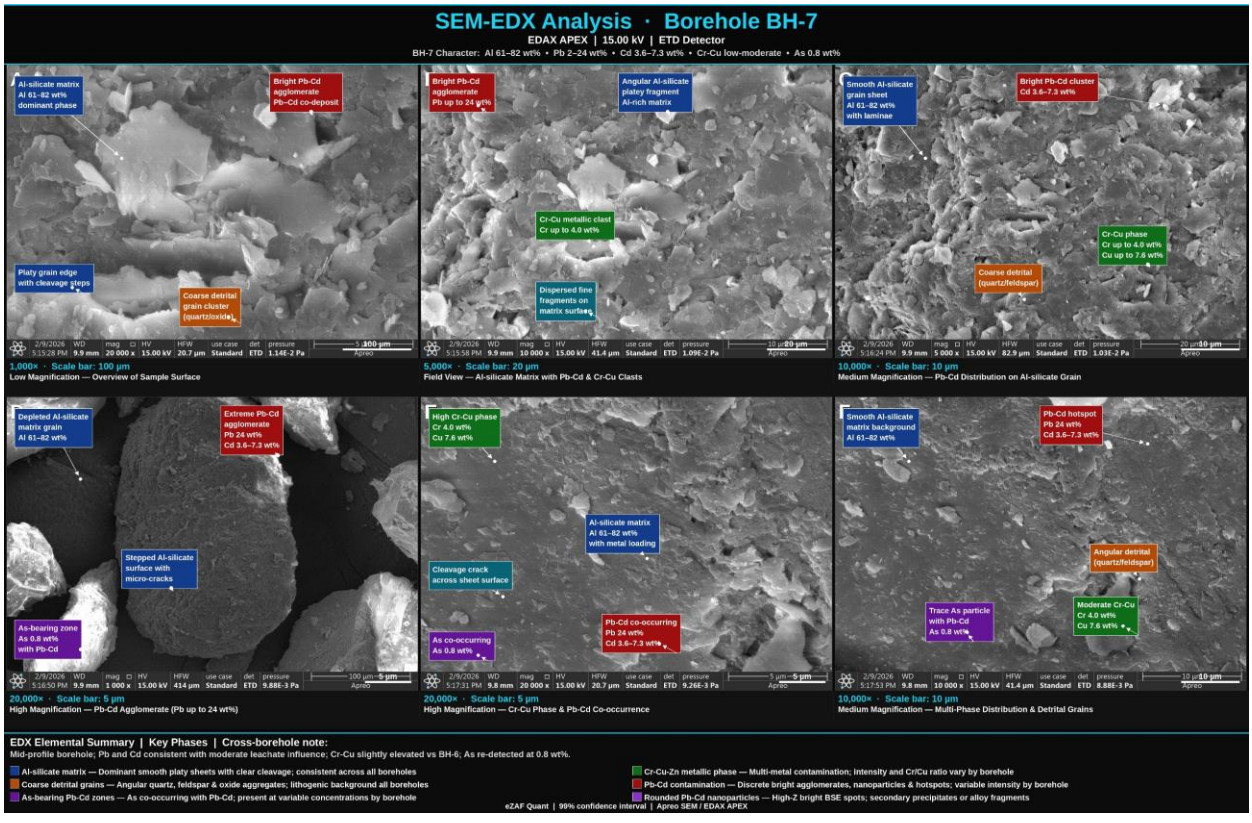
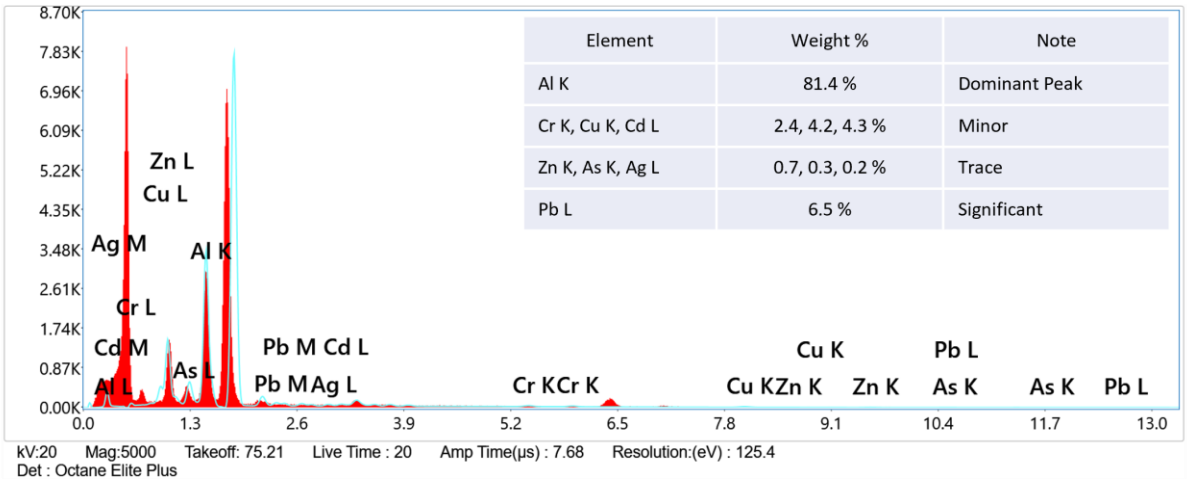


Figure 3.37 SEM observation of BH-7 soil sample

Bore Hole-8

BH-8 (Figure 3.38 – 3.39) reveals a pronounced re-intensification of contamination, with SEM micrographs showing a disrupted matrix surface covered by rough aggregates, fractured particles, fine debris, and several bright high-density inclusions.



Element	Weight %	Note
Al K	81.4 %	Dominant Peak
Cr K, Cu K, Cd L	2.4, 4.2, 4.3 %	Minor
Zn K, As K, Ag L	0.7, 0.3, 0.2 %	Trace
Pb L	6.5 %	Significant

Figure 3.38 EDS Spectrum — BH-8 (Highest Peak Area, EDAX APEX™)

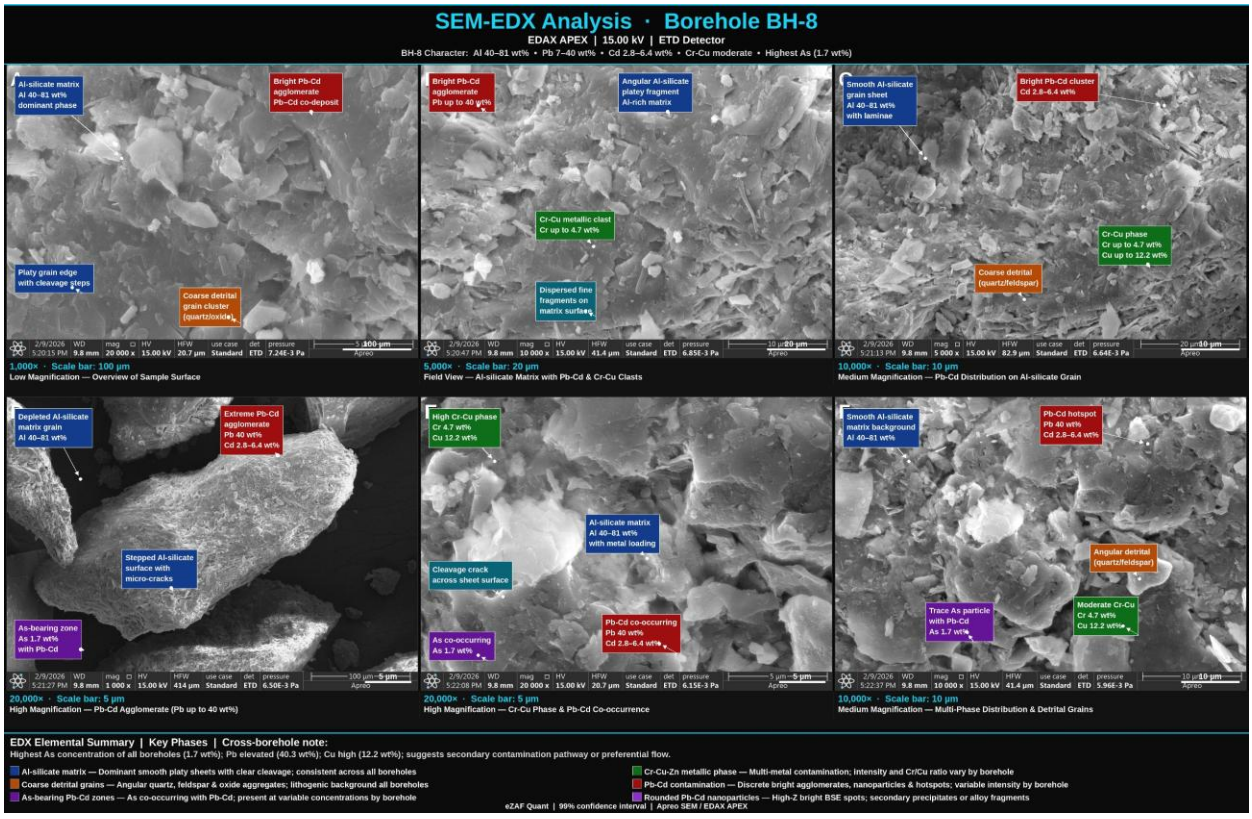


Figure 3.39 SEM observation of BH-8 soil sample

The Pb–Cd-rich particles appear as distinct bright clusters and agglomerates, indicating localized accumulation rather than uniform distribution over the matrix surface. Separate Cr–Cu-rich phases are also visible in different analytical areas, suggesting that Pb-rich and Cr–Cu-rich contaminants are not fully mixed but occur as independent metal-bearing particles within the sample. The presence of arsenic in selected zones further indicates association with these contaminant phases, although it appears less dominant than the major Pb, Cr, and Cu signatures. Compared with the more moderate pattern in BH-6, BH-8 shows stronger spatial variation, where heavy-metal-bearing particles seem to be trapped or accumulated within irregular surface pockets and fractured zones. Overall, BH-8 represents a distinctly impacted location, characterized by discrete Pb–Cd hotspots, elevated Cr–Cu phases, and uneven contaminant distribution within a visibly rough and fragmented microstructure.

Bore Hole-9

BH-9 (Figure 3.40 – 3.41) represents an anomalous distal contamination zone where the SEM–EDX evidence does not follow a simple decreasing trend with distance. The SEM micrographs show rough, fractured, and irregular particle surfaces with abundant fine debris, angular fragments, and several bright high-density inclusions attached to the matrix. Pb–Cd-rich clusters are clearly visible as localized agglomerates, indicating that lead-bearing contamination remains concentrated in selected hotspots rather than being uniformly distributed. At the same time, Cd-rich signatures occur even in areas where Pb is comparatively lower, suggesting possible differential mobility of cadmium and partial separation from Pb-dominant phases during leachate transport. Cu–Cr-bearing phases, with associated Zn and As in selected spots, are also observed as separate angular fragments, showing that multiple contaminant-bearing phases are present within the same borehole. The coexistence of relatively cleaner matrix areas with highly

contaminated particles indicates strong spatial contrast, likely controlled by preferential leachate pathways, localized trapping, or accumulation of dense metal-bearing particles.

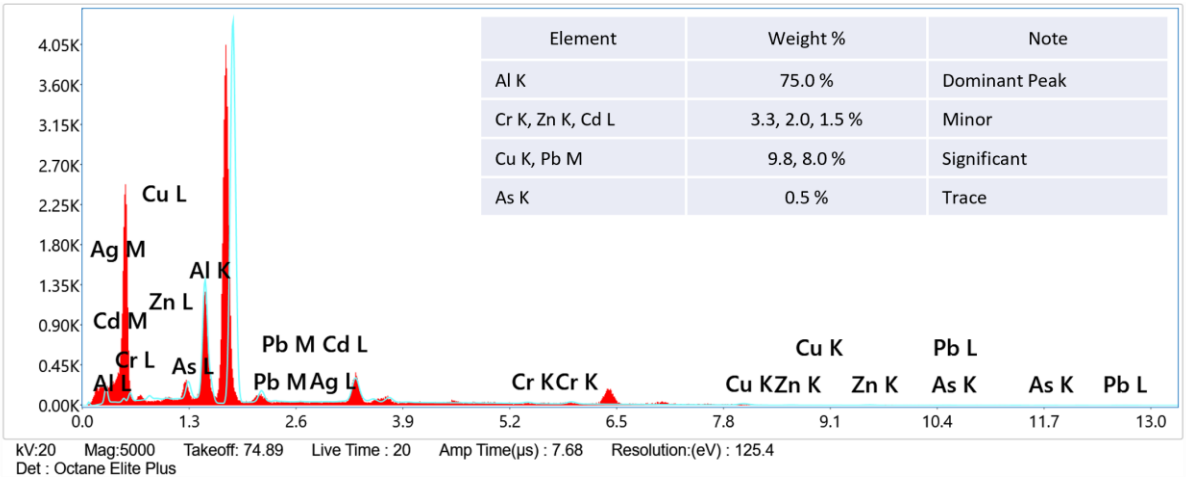


Figure 3.40 EDS Spectrum — BH-9 (Highest Peak Area, EDAX APEX™)

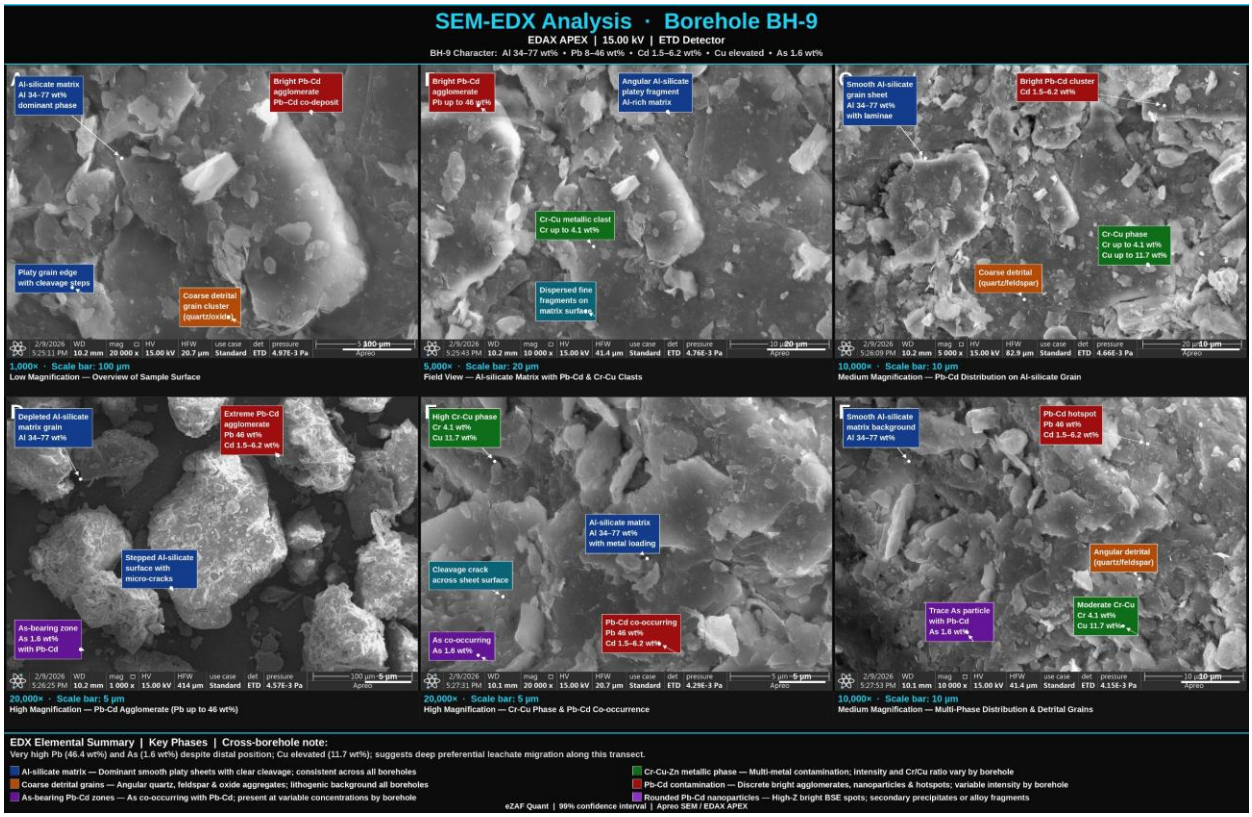
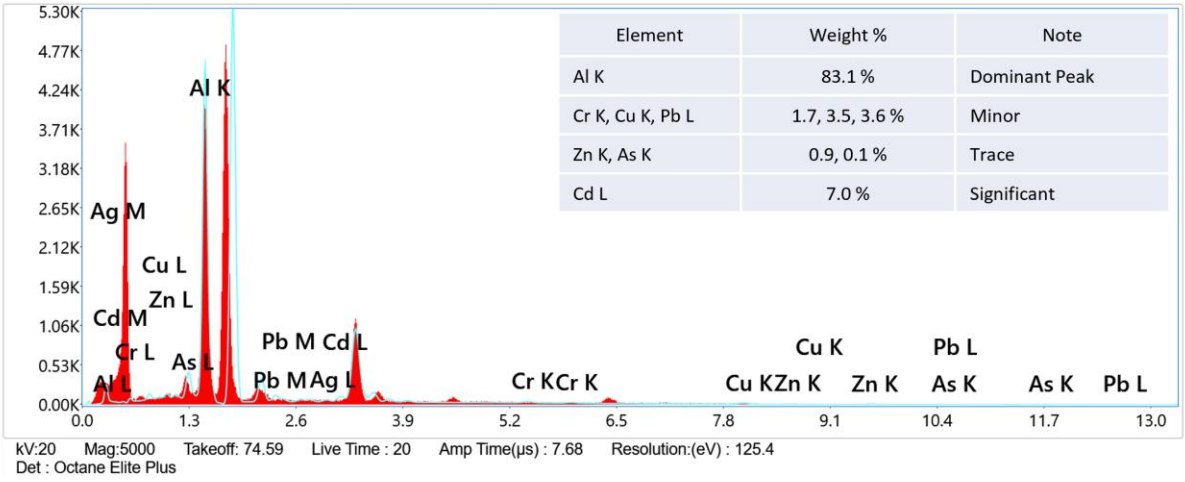


Figure 3.41 SEM observation of BH-9 soil sample

Overall, BH-9 reflects persistent and uneven metal contamination, marked by Pb–Cd hotspots, Cd-enriched zones, and Cu–Cr–Zn–As associations within a visibly disturbed microstructure.

Bore Hole-10

BH-10 (Figure 3.42 – 3.43) shows a more stable and background-approaching SEM–EDX pattern compared to the earlier boreholes, although complete absence of contamination is not indicated. The SEM micrographs show relatively smooth Al-rich matrix surfaces with fewer intense bright inclusions, scattered angular detrital particles, and limited fine particulate deposition. Pb–Cd-rich particles are still observed, but they appear as small, isolated clusters rather than large dominant hotspots, suggesting that Pb-bearing contamination has substantially reduced by this location. A notable feature of BH-10 is the persistent Cd signal across the analysed areas, indicating that cadmium remains comparatively mobile and continues to occur even where other metals have attenuated. Cr–Cu–Zn phases are weak and dispersed, showing limited influence of these contaminants at this borehole, while arsenic is restricted to trace-level association in selected zones. Overall, BH-10 represents the most attenuated end-member of the transect, where the matrix appears comparatively preserved, Pb, Cr, Cu, Zn and As signatures are reduced, but cadmium remains consistently present as the main residual contaminant.



Element	Weight %	Note
Al K	83.1 %	Dominant Peak
Cr K, Cu K, Pb L	1.7, 3.5, 3.6 %	Minor
Zn K, As K	0.9, 0.1 %	Trace
Cd L	7.0 %	Significant

Figure 3.42 EDS Spectrum — BH-10 (Highest Peak Area, EDAX APEX™)

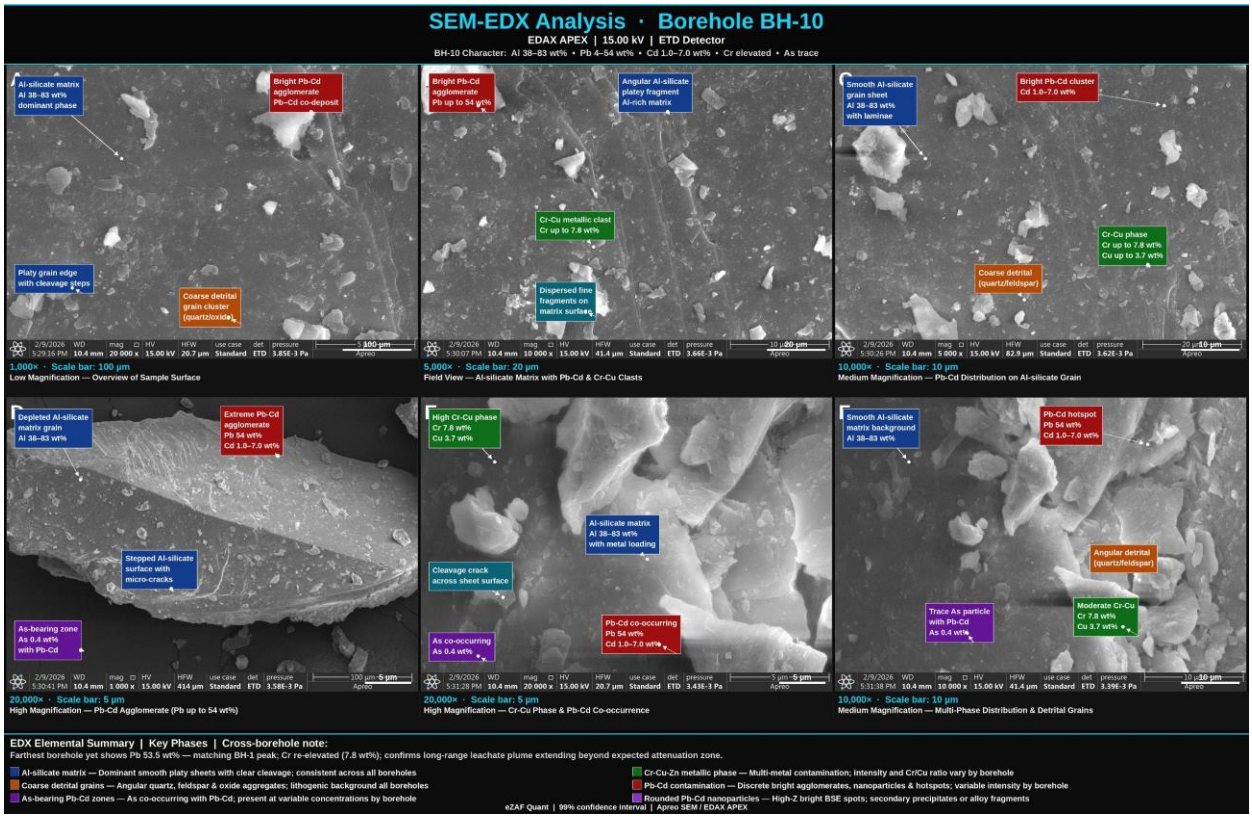


Figure 3.43 SEM observation of BH-10 soil sample

3.7 CONCLUSIONS

This chapter presented an integrated assessment of landfill leachate and leachate-impacted soil around the Bhalswa landfill, with emphasis on physicochemical characteristics, heavy metal contamination, pollution indices, geotechnical behaviour and SEM–EDX-based microstructural

interpretation. The chapter established the role of landfill leachate as a contaminant source and examined its possible influence on the surrounding subsurface soil through chemical, statistical, index-based and engineering-property evaluations.

The characterization of Bhalswa landfill leachate indicated that the leachate was **contaminated**. Although the pH of the leachate samples remained within the alkaline range, other parameters showed high pollution strength. Elevated **EC, TDS, TSS, chloride, nitrate, total hardness and COD** reflected high dissolved ionic load, suspended matter, salt enrichment and organic pollution. Heavy metal analysis showed that **Ni, Pb, Cr, Fe, Cd and Al** exceeded the adopted reference values, confirming toxic metal contamination in the leachate. Comparison with the **Solid Waste Management Rules, 2016** and **BIS IS 10500:2012** reference values further confirmed the contaminating potential of the leachate. The **Leachate Pollution Index values ranged from 14.89 to 16.71**, with an **average value of 15.93**, indicating measurable pollution potential. The LPI was mainly influenced by **COD, chloride, TDS, Cr, Pb, Ni, Cu and Fe**, confirming that Bhalswa landfill leachate can act as a significant source of contamination for surrounding soil and groundwater systems.

The leachate characterization confirmed that Bhalswa landfill acts as a major contamination source, as elevated physicochemical parameters and heavy metal concentrations were observed in the collected leachate samples. The LPI assessment further supported the pollution potential of the leachate and indicated its capacity to affect the surrounding subsurface environment.

The physicochemical assessment of soil samples showed clear variation across borehole locations and depths. Soil pH ranged from 6.30 to 8.22, indicating slightly acidic to alkaline subsurface conditions. EC, TDS, salinity and resistivity showed location-specific variation, with higher EC, TDS and salinity and lower resistivity at BH-6 at 9 m, indicating localized ionic enrichment. Alkalinity, chloride, calcium and total hardness also varied considerably, suggesting the presence

of soluble salts and leachate-associated ionic constituents within the soil matrix. These physicochemical variations support the interpretation that leachate interaction has influenced the chemical condition of the surrounding soil.

The heavy metal assessment identified **Zn, Pb, Mn and Cd** as the main contaminants in the leachate-impacted soil. **Zinc showed the highest frequency of exceedance**, crossing the Indian natural background value in **43 out of 50 samples**. **Lead exceeded** the background value in **30 samples**, while **Mn exceeded** it in **19 samples**. **Cadmium exceeded** the lower Indian soil limit in **10 samples** and the upper limit in **3 samples**. The maximum **Zn, Mn and Fe** concentrations were observed at **BH-6 at 12 m**, while the maximum **Pb and Cd** concentrations were recorded at **BH-5 at 9 m**. Based on exceedance frequency and comparison with adopted reference/background values, the contamination significance followed the order: **Zn > Pb > Mn > Cd > Fe > Cu > Cr > Ni > As > Co**. The spatial and vertical distribution showed that contamination was not uniform; instead, metal enrichment occurred in **localized zones**, particularly at **intermediate and deeper depths**. This pattern is important because Bhalswa landfill is an **unlined/non-engineered landfill**, allowing leachate to percolate downward and laterally through the soil profile.

The statistical and index-based assessments strengthened the interpretation of heavy metal contamination. Descriptive statistics showed wide concentration ranges for several metals, indicating **non-uniform enrichment** across boreholes and depths. Spearman correlation and regression analyses helped examine the influence of depth and distance on metal distribution. The **Geo-accumulation Index (Igeo)** indicated that **Cd and Pb reached Class 3**, representing **moderately to heavily contaminated** conditions, while **Zn reached Class 2**, indicating **moderate contamination**. **Mn reached Class 1**, reflecting **uncontaminated to moderately contaminated** status, whereas **Cu, As, Cr and Fe remained in Class 0**. The **Enrichment Factor (EF)** assessment showed stronger enrichment severity for **Cd, Pb, Zn and Mn**. **Cd recorded the highest mean EF value of 9.663**, with values reaching **66.013**, indicating **significant to**

extremely high enrichment in some samples. **Pb showed a mean EF of 5.477** and maximum EF of **42.105**, while **Zn showed a mean EF of 5.010**, with **all 50 samples exceeding the enrichment threshold**. **Mn showed moderate enrichment**, with a mean EF of **2.436**. In contrast, **Cu, As and Cr** showed deficiency to minimal enrichment. These results confirm that **Zn, Cd, Pb and Mn** are the main contamination and enrichment indicators in the study area, suggesting strong anthropogenic input associated with **landfill leachate, waste decomposition and localized subsurface migration**.

The geotechnical assessment demonstrated that leachate interaction has influenced the engineering behaviour of the soil. Although the soil profile was dominantly sandy, **fine-rich and plastic layers** were observed at deeper depths near the landfill, indicating localized alteration of the natural soil condition. Compaction characteristics were compared with **BH-10**, considered the least affected borehole, and the results showed that leachate-affected layers generally required **higher moisture for compaction** and developed **lower dry density**. This indicates reduced compactability due to **leachate migration, fines accumulation and moisture retention**. Permeability and shear strength behaviour also reflected this alteration. Fine-rich layers showed **reduced permeability and lower drainage capacity**, while shear strength behaviour shifted from predominantly frictional sandy response to more cohesive behaviour in affected layers. This was reflected by **higher cohesion and lower angle of internal friction** in contaminated/fine-rich zones. SEM–EDX analysis supported these findings by showing **microstructural variation and elemental evidence of contamination** within the soil matrix.

The soil quality assessment showed that the landfill-affected soil acts as an important pathway and intermediate sink for contaminant migration. The Igeo and EF assessments confirmed that the heavy metal contamination in soil is not only due to natural background variation but also reflects anthropogenic enrichment associated with landfill leachate and waste decomposition. Thus, the

soil contamination results provide an important link between the polluted leachate source and possible groundwater quality deterioration.

In conclusion, Chapter 3 demonstrates that Bhalswa landfill leachate has influenced the physicochemical, heavy metal and geotechnical characteristics of nearby soil. The combined evidence from leachate characterization, LPI, soil chemistry, heavy metal distribution, statistical analysis, Igeo, EF, geotechnical testing and SEM–EDX confirms that the surrounding soil has undergone ionic enrichment, metal accumulation and engineering-property modification due to leachate interaction. These findings provide an important basis for understanding contaminant migration from the landfill and for interpreting groundwater quality variations discussed in **Chapter 4**.

CHAPTER 4: HYDROGEOCHEMICAL ANALYSIS OF GROUNDWATER CONTAMINATION DUE TO LANDFILL LEACHATE

4.1 INTRODUCTION

Water bodies cover around 71% of the Earth's surface, totaling 1,386 million km³, of which only 0.5% is accessible freshwater (WMO, 2022). Water scarcity is a serious problem; despite its abundance, approximately half the World's population will face severe shortages at least part of the year (UNESCO, 2024). The global water demand is expected to rise by 20% to 30% by 2050 due to population growth and economic expansion, further straining these finite resources (UN, 2018). Groundwater (GW) provides more freshwater than lakes and streams, but is hard to collect, slow to recover, and easily contaminated. GW supplies 25% of irrigation water and half of residential freshwater (UNESCO, 2024). The composition of water plays an important role in determining its quality for consumption (Raheja et al., 2022). The physical and chemical properties of groundwater have been degraded by overexploitation in regions with high climate variability. Globally, groundwater contamination is a significant issue with major impacts on public health (Li et al., 2021). Groundwater constitutes approximately 98% of the global freshwater reserves, which are extensively used for household, agricultural, and industrial purposes (WWAP, 2022).

Total water utilization in India is approximately 1,123 billion cubic meters, of which 690 billion cubic meters are managed by surface water resources and the remaining 433 billion cubic meters by groundwater (Ministry of Water Resources, 2011). Groundwater requires more energy to bring the water to the surface. Generally, in groundwater, total dissolved solids (TDS), total Hardness, and Total Alkalinity were higher than the permissible levels, and studies found that urban areas have higher loading values for Fe and Cl- than semi-urban areas, due to different types of contamination (Bose et al., 2024). Groundwater quality is depleted by several sources, such as effluent from industries and the percolation of chemicals used in agriculture, which are well-

known sources. In urban regions, sources are even more complex due to rapid urbanization. Solid waste management is a significant challenge in urban regions, negatively impacting air and surface water quality and creating aesthetic issues. Despite these known issues, the regions adjacent to solid waste dumpsites also face additional problems, including contamination of surface water and deterioration of groundwater quality.

India, one of the World's second-most populous countries, has a water demand that consistently exceeds supply. Total water utilization in India is approximately 1,123 billion cubic meters, of which 690 billion cubic meters are managed by surface water resources and the remaining 433 billion cubic meters by groundwater (Ministry of Water Resources, 2011). Groundwater requires more energy to bring the water to the surface. Generally, in groundwater, total dissolved solids (TDS), total hardness, and total alkalinity were higher than the permissible levels, and studies found that urban areas have higher loading values for Fe and Cl^- than semi-urban areas, due to different types of contamination (Bose et al., 2024).

Groundwater quality varies in both horizontal and vertical directions due to changes in aquifer boundaries or variations within a single geological structure (Qiao et al., 2025). The groundwater capacity of the aquifer type varies significantly from place to place and from basin to basin (Falowo et al., 2023). The use of groundwater has significantly boosted agricultural activities, enhancing the livelihoods and food security of billions of farmers and consumers worldwide. Excessive groundwater usage leads to several social, economic, and environmental consequences. (Parakh, 2022).

This chapter describes the quality of groundwater from various water sources, including boreholes, handpumps, submersible, and surface water samples collected near the Bhalswa dump. The findings help identify contamination intensity, affected zones, and possible pathways by which

leachate influences groundwater quality. These results establish a foundation for understanding the impact of leachate on groundwater quality, which is addressed in **Chapter 3.**

4.2 MATERIALS AND METHODS

4.2.1 STUDY AREA AND SAMPLING LOCATIONS

The Bhalswa landfill and its surrounding areas were selected as the study area (28.7395° N, 77.1573° E), as illustrated in Figure 4.1.

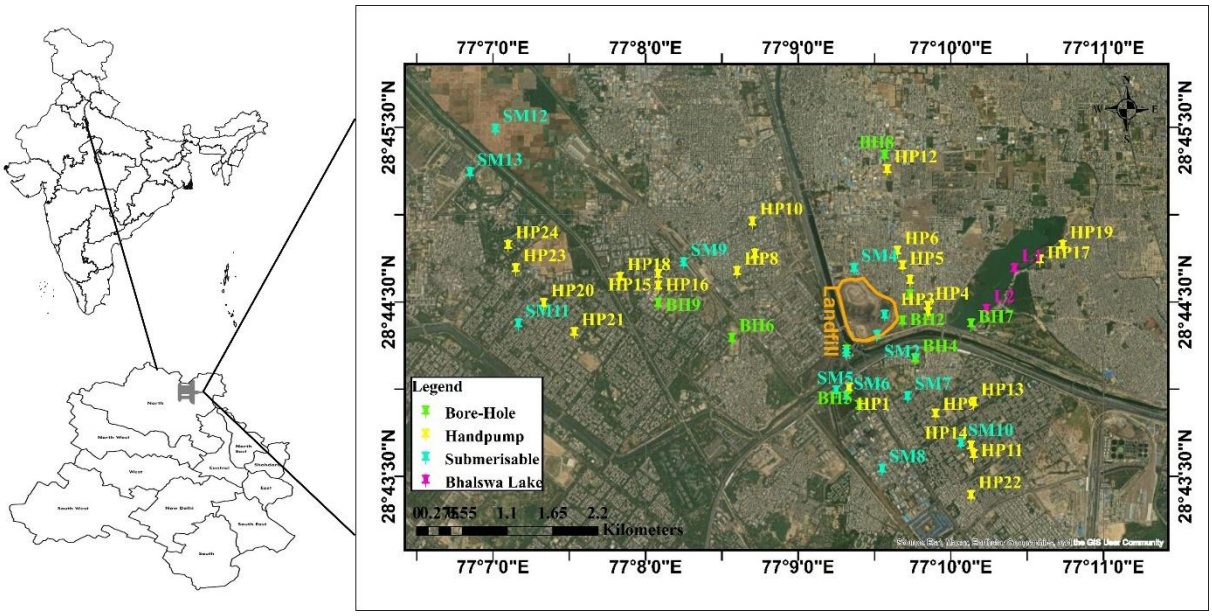


Figure 4.1 Groundwater Sampling location in the vicinity of Bhalswa landfill, Delhi.

The study area falls under the jurisdiction of the Municipal Corporation of Delhi (MCD). The study region experiences a range of extreme climatic conditions. The climatic conditions in the study area range from 25°C to 45°C during the summer, and from 5°C to 25°C during the winter. The Bhalswa landfill gets approximately 2500 metric tons of garbage daily. Since 1994, this dumpsite has received 2,500 metric tons of solid waste daily (CPCB, 2024; Bharadwaj et al.,

2025a). The waste collected at the landfill region consists of unsegregated municipal waste, including plastics, metals, inert debris, and both biodegradable and non-biodegradable wastes. The study area is characterized by hard rock terrain with extensive alluvial deposits. The landfill is covered with quaternary deposits, which vary in thickness from one location to another. The sediments consist of fine to medium sand, often referred to as Yamuna sand. The topmost soil layer identified is sand and gravel, extending to a depth of up to 40 meters. The layer is followed by alluvium, composed primarily of compact clay, silt, and kankar nodules (CPCB, 2024).

4.2.2 GROUNDWATER AND SURFACE WATER SAMPLING

Groundwater samples (Figure 4.2) were systematically collected from various locations across the Bhalswa landfill, extending towards the residential zones, to evaluate spatial variability in groundwater quality. A total of 46 groundwater samples were collected, comprising 9 borehole samples, 24 hand-pump samples, and 13 submersible-pump samples, as shown in Figure 4.1. Groundwater samples were preserved in pre-cleaned polyethylene bottles in accordance with standard procedures IS 3025 (Part 1): 1987. Sources from different depths, hydrogeological conditions, and anthropogenic influence across the study area were chosen. Boreholes and submersible pumps revealed deeper aquifer conditions, while hand pumps revealed shallow groundwater properties and surface contamination risk. Surface water samples were obtained from Bhalswa Lake: one upstream and one downstream, as depicted in Figure 4.1.

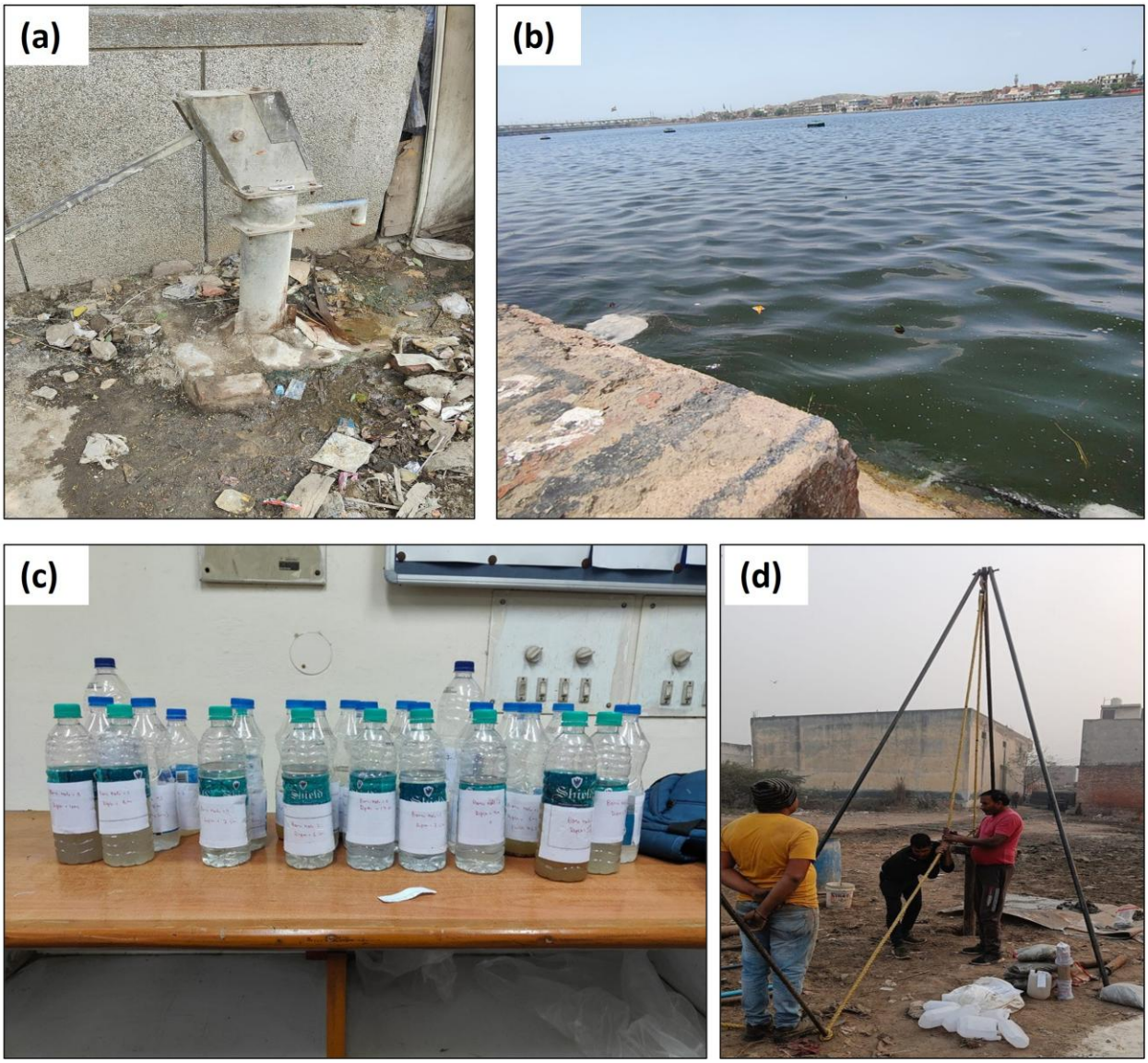


Figure 4.2 (a) GW sample collection from Handpump (b) Bhalswa lake (c) GW samples from different sources (d) GW sample collection from Borehole location in the vicinity of Bhalswa landfill, Delhi.

Multi Parameter Water Quality Meter

The pH, electrical conductivity (EC), salinity, and resistivity were assessed with a calibrated multi-parameter water quality meter (HACH/Thermo).

pH was determined as per IS 3025 (Part 11): 2022 and EC per IS 3025 (Part 14): 2013.

Chloride

The chloride in the water sample is determined using titration with silver nitrate solution (IS 3025 Part 32). A measured sample volume was titrated against standard $AgNO_3$ solution using potassium chromate (K_2 Cr O_4) as an indicator. The endpoint was defined by the formation of a brick-red Ag_2CrO_4 precipitate (Argentometric Method).

Alkalinity

The alkalinity of the water sample is determined by titrating H_2SO_4 to the phenolphthalein endpoint, followed by the methyl orange endpoint. The titration was conducted using a standardized 0.02 N H_2SO_4 solution until the endpoint was achieved, indicated by a pinkish color (IS 3025 Part 23).

Total hardness – Calcium Ion and Magnesium Ion

Total hardness was determined by the standard EDTA Titrimetric Method. The total hardness (TH) was calculated based on the ability of ethylenediaminetetraacetic acid (EDTA) to form stable complexes with Ca^{2+} and Mg^{2+} ions. The sample was buffered to pH ≈10 using an NH_4Cl–NH_4OH buffer; Eriochrome Black T indicator was added (producing a wine-red colour), and the sample was titrated with standard 0.01 M EDTA until the colour changed sharply to blue, indicating complete sequestration of Ca^{2+} and Mg^{2+}. (IS 3025 (Part 21): 2009). Calcium was determined separately by titrating a fixed sample volume with EDTA using murexide indicator. The total harness is calculated using Equation 4.1.

$$\textbf{Total Hardness (as } CaCo_3) \frac{\textbf{mg}}{\textbf{L}} = \left(\frac{\textbf{Volume of EDTA} \times \textbf{Molarity of EDTA} \times \textbf{Molecular weight of } CaCo_3}{\textbf{Volume of sample}}\right) \quad \textbf{(4.1)}$$

- Where, Molecular weight of $CaCO_3$ = 100 g/mol

Then using Equation 4.2, we have calculated Magnesium hardness (IS 3025 Part 46).

$$\textbf{Total Hardness (as } CaCo_3) = \textbf{Calcium hardness (as } CaCo^3) + \textbf{Magnesium hardness (as } CaCo_3) \quad \textbf{(4.2)}$$

Heavy metal Analysis

Groundwater and surface water samples for heavy metal analysis is filtered through a 0.45 µm membrane filter (IS 3025 (Part 2): 2004). The filter is then acidified using nitric acid. The sample is then stored in an ice box (4˚C). The heavy metal analysis is done using Microwave Plasma Atomic Emission Spectroscopy (MP-AES). These instruments work on the principle of microwave-induced plasma to atomize and excite the sample. Atomic emission can also be used to identify the elements present in the sample.

Dissolved oxy gen and Salinity for surface water sample

Dissolved Oxygen in the freshwater samples of Bhalswa Lake was determined using the titration method (IS 3025 Part 38), in which 2 mL of manganese sulphate and an additional 2mL of acetic acid were used to form a brown precipitate. 1 mL of sulfuric acid was added to turn the solution transparent yellow. Then the sample is titrated with sodium thiosulfate ($Na_2S_2O_3$) until the solution turns colorless.

4.2.3 WEIGHTED ARITHMETIC WATER QUALITY INDEX (WA-WQI)

The water quality index (WQI) is a quantitative measure used to assess the suitability of water for several applications, including drinking, industrial purposes, and supporting aquatic life. The WA-WQI approach directly supports monitoring of SDG 6, Targets 6.3 (improving water quality) and Target 6.1 (access to safe drinking water) (Ngwenya et al., 2025). The method most frequently utilized to analyse WQI is the weighted arithmetic method. It categorizes the purity level of water based on frequently measured water parameters (Noori et al. 2020). The steps in calculating the Weighted Arithmetic Water Quality Index as given by researchers are as follows (Balan et al. 2012; Krishan et al. 2022).

STEP 1: The quality rating scale (Q) for each parameter is calculated using the formula

$$Q_i = [\frac{V_a - V_i}{V_s - V_i}] \times 100 \quad (4.3)$$

Where, V_a = Actual value of water quality parameter obtained from the data

V_i = Ideal value of water quality parameter (pH 7, all other parameters are zero)

V_s = Recommended standard value of the parameter

STEP 2: The proportionality constant K is calculated as

$$K = \frac{1}{\Sigma(\frac{1}{V_S})} \quad (4.4)$$

STEP 3: The unit weight for each parameter is calculated using the equation

$$W_i = \frac{K}{V_S} \quad (4.5)$$

STEP 4: The Water Quality Index is calculated as

$$WQI = \frac{\sum Q_i W_i}{\sum W_i} \quad (4.6)$$

4.3 RESULTS AND DISCUSSION OF GROUNDWATER QUALITY

4.3.1 PHYSICOCHEMICAL ANALYSIS FOR BOREHOLE SAMPLES

The Total Dissolved Solids (TDS) concentration ranged between the IS10500:2012 permissible concentrations, with a range of 318 to 470 mg/L (Figure 4.2). The total hardness concentration in the collected samples was found to be extremely higher (~940 mg/L) than the IS10500: 2012 permissible limits (200 mg/L), which is possibly due to the mixture of construction debris in the solid waste landfill. The calcium (88.9 to 636 mg/L) and magnesium (114 to 805.8 mg/L) (Figure

4.2) were found to be higher than the Indian permissible limits for both (100 and 200 mg/L) in the majority of the samples. A higher level of hardness has several impacts, including gastrointestinal irritation in humans upon consumption and plumbing damage in pipes during transport. The effect of a solid waste landfill on groundwater quality is further confirmed by total alkalinity and Chloride concentrations, which exceed the permissible limits as per IS standards. In BH 1, the chloride concentration exceeded 1000 mg/L due to the proximity of the Bore Hole to the landfill. Chloride ions are highly mobile and non-reactive, indicating that leachate from the landfill is affecting groundwater quality.

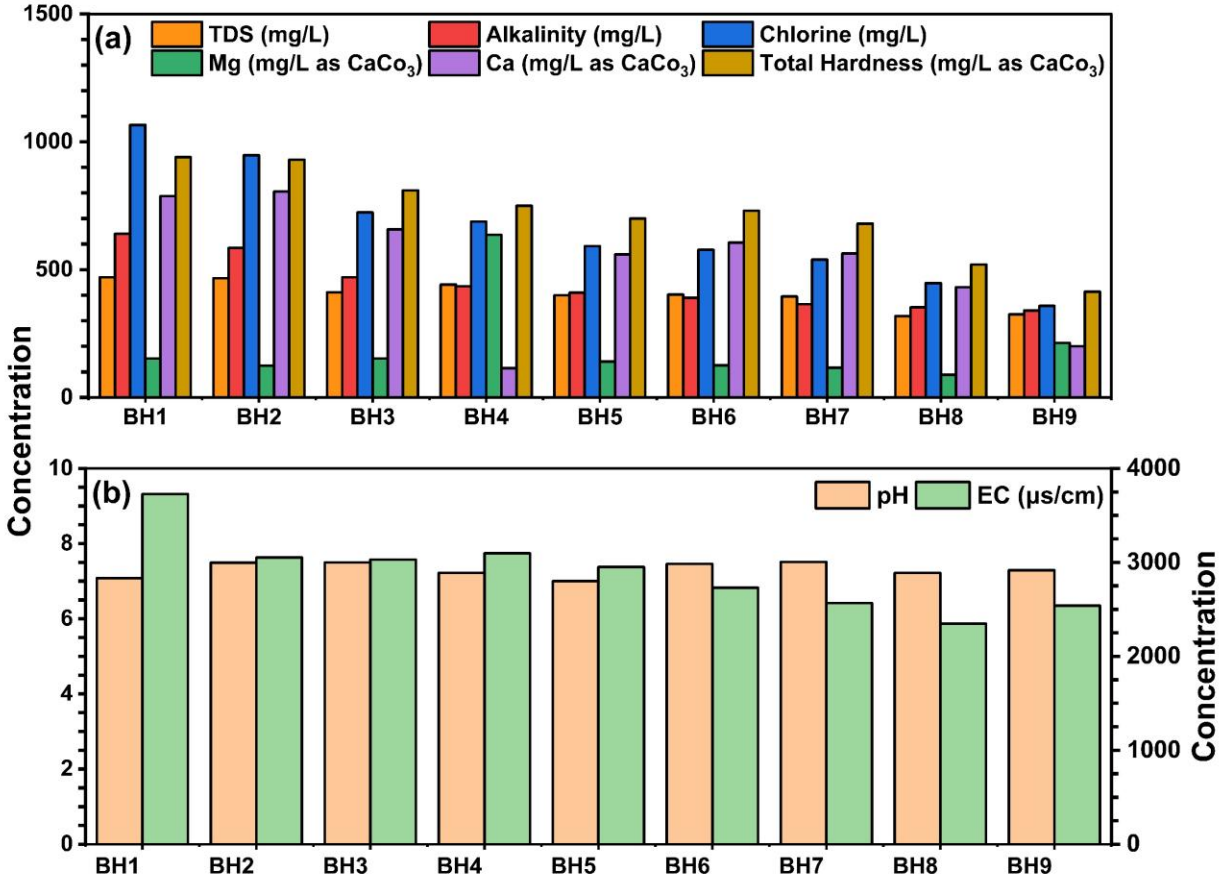


Figure 4.3 Physicochemical characteristics of borehole (BH1-BH9) groundwater samples of (a) EC, TDS, Major Ions, Hardness, and (b) pH.

The results revealed that the groundwater quality in the region is depleted due to active leachate migration, localized severe contamination, and long-term hydrochemical evolution in the

groundwater due to landfill contamination. Table 4.1 summarizes the physicochemical data of all borehole samples. The Electrical Conductivity EC concentration in the collected samples ranged between 2347 µs/cm and 3727.5 µs/cm (Figure 4.2), which is typically higher than the normal groundwater (IS 10500:2012). The increase in concentration of EC in the groundwater samples is due to the massive influx of higher amounts of dissolved organic ions from the landfill in the aquifer. Leaching of soluble salts and metallic ions is commonly found in heterogeneous urban waste decomposition. Potassium (K) is also exceeding permissible limit as given in Table 4.2.

Table 4.1 Physicochemical analysis of GW samples from borehole

Source	BH1	BH2	BH3	BH4	BH5	BH6	BH7	BH8	BH9	BIS 10500:2012 (Acceptable Limit)	BIS 10500:2012 (Permissible Limit)
pH	7.1	7.5	7.5	7.2	7.0	7.5	7.5	7.2	7.3	6.5–8.5	No relaxation
EC (µs/cm)	3727.5	3051.3	3030.0	3096.5	2950.0	2730.0	2567.0	2347.0	2540.0	Not	Not specified
TDS (mg/L)	470.0	466.5	411.2	441.7	400.0	402.1	394.8	318.1	324.7	500	2000
Alkalinity (mg/L)	640.0	585.0	470.0	435.0	410.0	390.0	365.0	352.0	340.0	200	600
Chloride (mg/L)	1066.2	947.9	723.5	687.5	592.0	577.5	539.0	447.0	358.0	250	1000
Magnesium (mg/L as $CaCo_3$)	152.3	124.2	152.3	636.0	140.3	125.0	116.3	88.9	213.5	30	100
Calcium (mg/L as $CaCo_3$)	787.7	805.8	657.7	114.0	559.7	605.5	563.7	431.1	200.0	75	200
Total Hardness (mg/L as $CaCo_3$)	940.0	930.0	810.0	750.0	700.0	730.5	680.0	520.0	413.5	200	600

4.3.2 HEAVY METAL CONTAMINATION FOR BOREHOLE SAMPLES

The heavy metal contamination analysis of groundwater revealed the presence of toxic heavy metals as shown in Table 4.2.

The overall order of heavy metal (Table 4.2) contamination by concentration was Fe > Al > Cd > Mn > Ni > Cr > Cu > As > Pb > Zn.

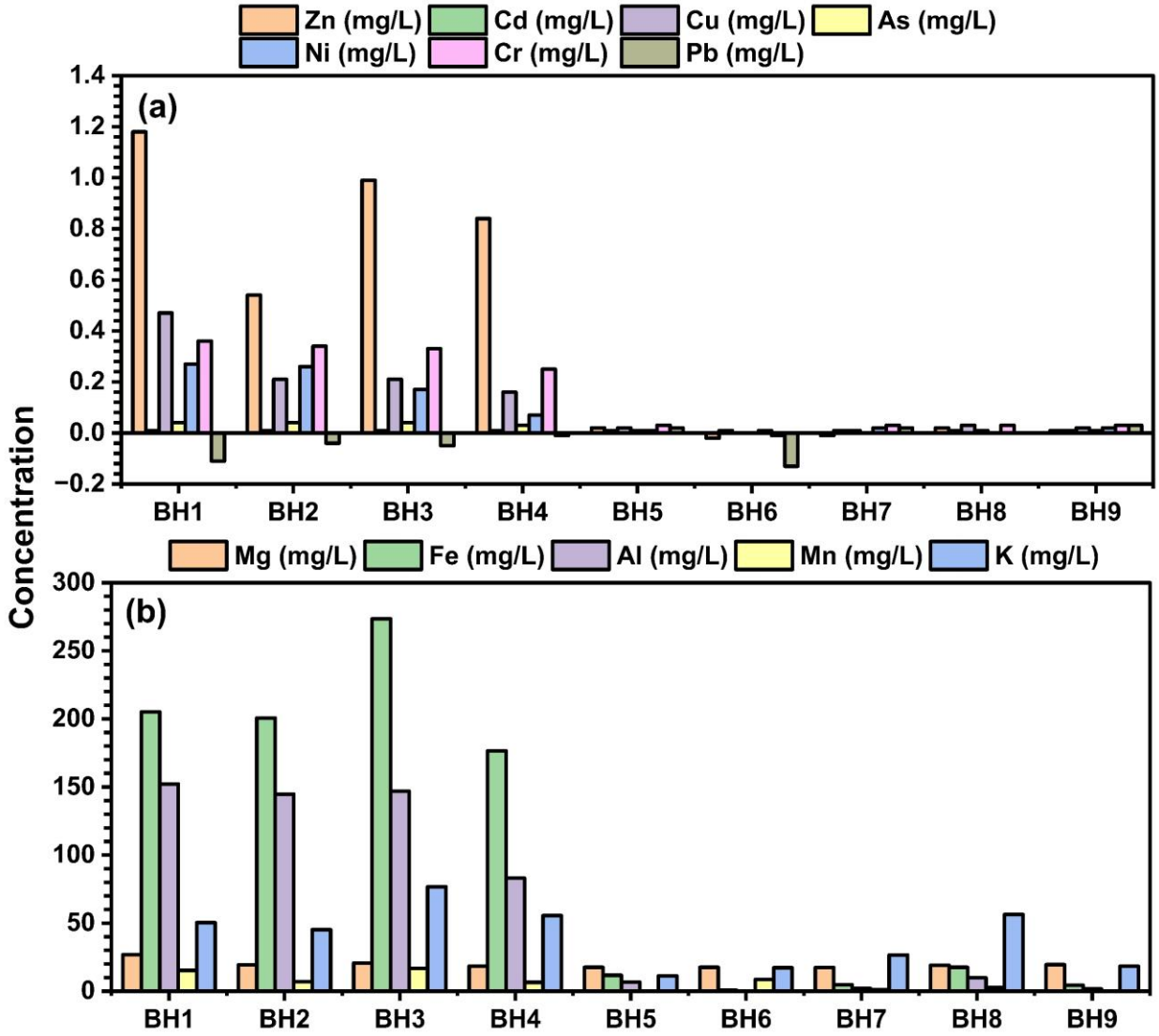


Figure 4.4 Heavy metal and major element concentrations in groundwater samples.

The Fe concentration in BH1 to BH4 (176.58–273.6 mg/L) is far exceeding the IS 10500:2012 standards (0.3 mg/L) (Figure 4.3). The higher concentration is due to the reduction of iron-bearing minerals at the soil surface under aerobic conditions created by organic-rich leachate. The maximum concentration of Al was found in Bore Hole 1(BH1), which is the nearest BH to the

Table 4.2 Heavy metal analysis of GW samples from borehole

Samples	BH1	BH2	BH3	BH4	BH5	BH6	BH7	BH8	BH9	BIS 10500:2012 (Acceptable Limit)	BIS 10500:2012 (Permissible Limit)
Zn (mg/L)	1.18	0.54	0.99	0.84	0.02	-0.02	-0.01	0.02	0.01	5	15
Cd (mg/L)	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.003	No relaxation
Cu (mg/L)	0.47	0.21	0.21	0.16	0.02	0	0.01	0.03	0.02	0.05	1.5
As (mg/L)	0.04	0.04	0.04	0.03	0.01	0	0	0.01	0.01	0.01	0.05
Mg (mg/L)	26.73	19.34	20.72	18.35	17.55	17.55	17.39	19.07	19.52	30	100
Ni (mg/L)	0.27	0.26	0.17	0.07	0.01	0.01	0.02	0	0.02	0.02	No relaxation
Fe (mg/L)	205.16	200.71	273.6	176.58	11.73	0.66	4.69	17.56	4.48	0.3	No relaxation
Al (mg/L)	152.62	144.78	146.93	83.03	6.66	0.19	1.95	9.86	1.63	0.03	0.2
Pb (mg/L)	-0.11	-0.04	-0.05	-0.01	0.02	-0.13	0.02	0	0.03	0.01	No relaxation
Mn (mg/L)	15.26	6.99	16.79	6.49	0.2	8.61	1.17	2.75	0.2	0.1	0.3
Cr (mg/L)	0.36	0.34	0.33	0.25	0.03	-0.01	0.03	0.03	0.03	0.05	No relaxation
K (mg/L)	50.4	45.19	76.72	55.59	11.31	17.23	26.59	56.31	18.27		<12 (WHO guidelines for driniking water)

(Yellow colour highlights the elements which are above the limits).

landfill site. The concentration observed is 152.26 mg/L, which is much higher than the Indian permissible limits (0.03 to 0.2 mg/L) (Figure 4.3). The manganese values were also found to be higher than the limits. Heavy metal concentrations, such as nickel and Chromium, in BH1 and BH4 are also found to exceed permissible limits.

The analysis revealed that the heavy metals such as Cd, Fe, Al, Mn, Cr, Ni, Cd, Cu, As and K exceed the limits, but Zn, and Mg are within the safe limits in the study area. BH1-BH4 show significantly higher contamination than BH5-BH9, showing the effect of leachate from landfills on the groundwater quality. Fe and Al are at critically dangerous levels requiring immediate attention, and Cd uniformly exceeds limits across all boreholes.

4.3.3 IDENTIFICATION OF CONTAMINATION IN HAND PUMP WATER SAMPLES

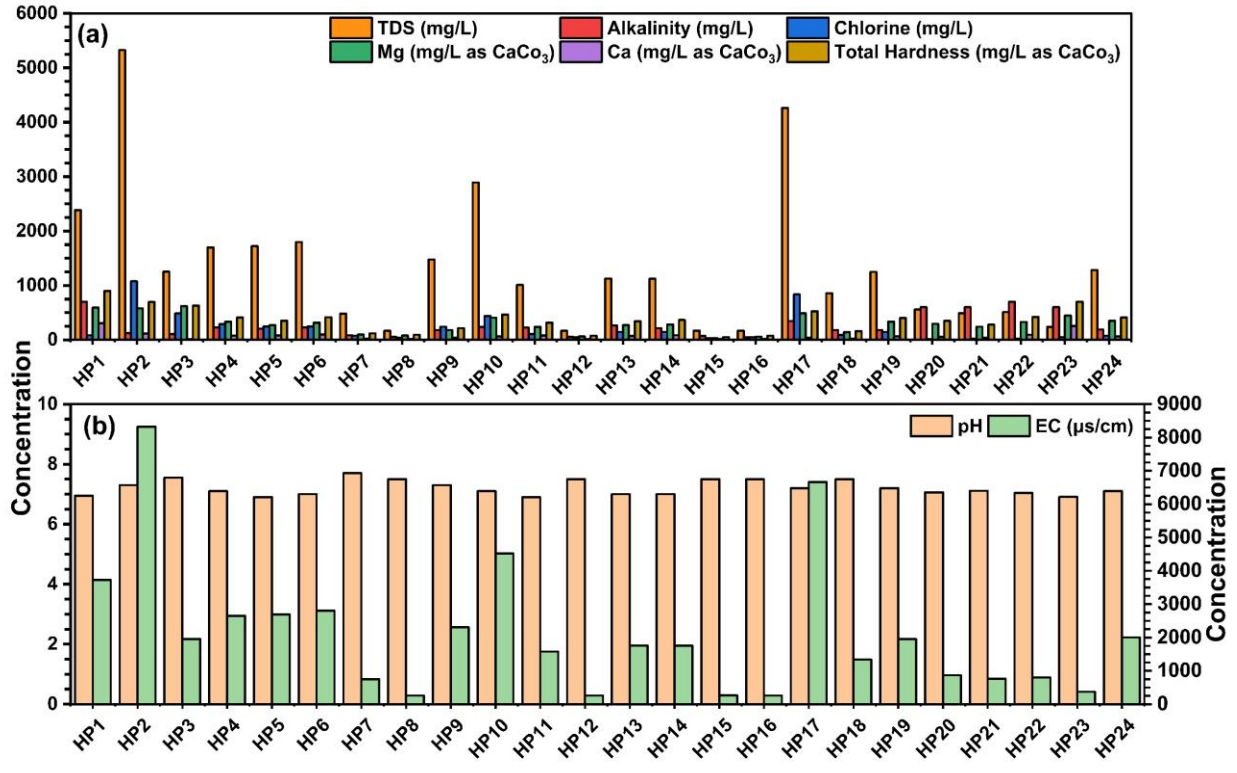


Figure 4.5 Physicochemical characteristics of hand pump (HP1-HP24) groundwater samples of (a) TDS, Major Ions, Hardness, and (b) EC, pH.

The EC concentration in the handpump water samples ranged from 259.7 to 8325 µs/cm (Figure 4.5); similarly, the TDS concentration was also observed to exceed the permissible limits, reaching a maximum of 5328mg/L. Hardness concentration observed in the handpump is higher (900 mg/L) (Figure 4.4), similar to the borehole concentration (940 mg/L). The maximum chloride concentration was 1076.18 mg/L, and the alkalinity was 700 mg/L (Figure 4.5). Table 4.3 describes the physicochemical analysis of handpump GW samples.

Table 4.3Physicochemical analysis of GW samples from handpumps

Source	pH	EC (µs/cm)	TDS (mg/L)	Alkalinity (mg/L)	Chloride (mg/L)	Magnesium (mg/L)	Calcium (mg/L)	Total Hardness (mg/L)
Hand Pump 1	6.95	3727	2385.3	700	85.2	591.4	308.6	900
Hand Pump 2	7.3	8325	5328	125	1076.18	579	116	695
Hand Pump 3	7.55	1956	1251.8	105	485	618.7	11.31	630
Hand Pump 4	7.1	2651	1696.6	230	290.319	334	76	410
Hand Pump 5	6.9	2694	1724.2	205	245.27	268	82	350
Hand Pump 6	7	2808	1797.1	230	245.27	317	98	415
Hand Pump 7	7.7	751	480.6	85	75.0825	96	24	120
Hand Pump 8	7.5	259.7	166.2	55	35.0385	79	16	95
Hand Pump 9	7.3	2306	1475.8	180	240.264	179	36	215
Hand Pump 10	7.1	4520	2892.8	235	438.982	405	60	465
Hand Pump 11	6.9	1577	1009.3	225	105.116	237	78	315
Hand Pump 12	7.5	260.7	166.8	55	45.0495	63	12	75
Hand Pump 13	7	1760	1126.4	265	145.16	270	70	340
Hand Pump 14	7	1756	1123.8	215	145.16	281	84	365
Hand Pump 15	7.5	262.1	167.7	70	30.033	34	16	50
Hand Pump 16	7.5	261.3	167.2	45	50.055	59	16	75
Hand Pump 17	7.2	6663	4264.3	345	833.416	489	36	525
Hand Pump 18	7.5	1339	857	180	90.099	140	20	160
Hand Pump 19	7.2	1952	1249.3	180	145.16	334	66	400
Hand Pump 20	7.06	869.5	556.5	600	14.2	293.9	56.11	350
Hand Pump 21	7.11	765.4	490.9	600	17.75	239.9	40.08	280
Hand Pump 22	7.04	799.8	511.9	700	17.75	327.8	92.18	420
Hand Pump 23	6.91	372	238.1	600	49.7	447.5	252.5	700
Hand Pump 24	7.1	2004	1282.6	190	75.0825	350	60	410
BIS 10500:2012 (Acceptable Limit)	6.5–8.5	Not specified	500	200	250	30	75	200
BIS 10500:2012 (Permissible Limit)	No relaxation	Not specified	2000	600	1000	100	200	600

Alkalinity, chloride, magnesium, calcium, and total hardness also showed higher concentrations in many samples than the desirable and permissible drinking water limits.

The increased concentrations of these parameters suggest a strong interaction between groundwater and rock, especially in areas dominated by carbonate and chloride-rich minerals. High hardness values indicate a significant presence of calcium and magnesium ions, which may result from the dissolution of limestone, dolomite, or other mineral formations present in the subsurface.

4.3.4 IDENTIFICATION OF CONTAMINATION IN SUBMERSIBLE PUMP WATER SAMPLES

All collected submersible (SM) samples physicochemical properties are illustrated in Table 4.4. The EC concentration exceeded the permissible limits in more than 60% of the collected samples. The values ranged from 260.3 µs/cm to 6610 µs/cm (Figure 4.5), where the permissible limit is 1500 µs/cm. The maximum concentration of TDS was found to be 4320 mg/L (Figure 4.5). Out of 13 samples collected from the submersible pumps, only 4 pump concentrations were found to be in the permissible limit, and 3 pump concentrations were in the acceptable range, with the remaining pump concentrations found to be more than the acceptable range and permissible range. The total hardness concentration is found to be higher in 4 samples with a maximum concentration of 1210 mg/L (Figure 4.5). Similar to the hardness, the calcium and magnesium are also found to be higher in the same samples, where the hardness is driven by a higher concentration of calcium and magnesium salts in the water sample. The alkalinity was found to be lower in only three submersible water samples; in the other samples, the concentration is either slightly higher or much higher than the standards.

Table 4.4 Physicochemical analysis of GW samples from submersible pump samples

Source	pH	EC (µs/cm)	TDS (mg/L)	Alkalinity (mg/L)	Chlorine (mg/L)	Magnesium (mg/L)	Calcium (mg/L)	Total Hardness (mg/L)
Submersible 1	7.2	5080	3251	210	495.545	427	98	525
Submersible 2	6.81	6610	4230	1700	124.3	777.1	432.9	1210
Submersible 3	6.87	3557	2276	1200	63.9	593.3	336.7	930
Submersible 4	6.91	3354	2147	1100	53.25	497.6	212.4	710
Submersible 5	7.4	2375	1520	210	200.22	289	56	345
Submersible 6	7.6	1868	1196	250	145.16	199	36	235
Submersible 7	7.3	1455	931	170	110.121	251	54	305
Submersible 8	7.1	5914	3785	200	545.6	705	160	865
Submersible 9	7.5	260.3	167	45	25.0275	72	18	90
Submersible 10	7.2	4130	2634	335	545.6	385	80	465
Submersible 11	7.6	452.5	290	70	55.0605	104	26	130
Submersible 12	7.17	341.6	219	700	10.65	69.9	72.14	290
Submersible 13	6.74	275.4	176	300	10.64	89.7	20.04	190
BIS 10500:2012 (Acceptable Limit)	6.5–8.5	Not specified	500	200	250	30	75	200
BIS 10500:2012 (Permissible Limit)	No relaxation	Not specified	2000	600	1000	100	200	600

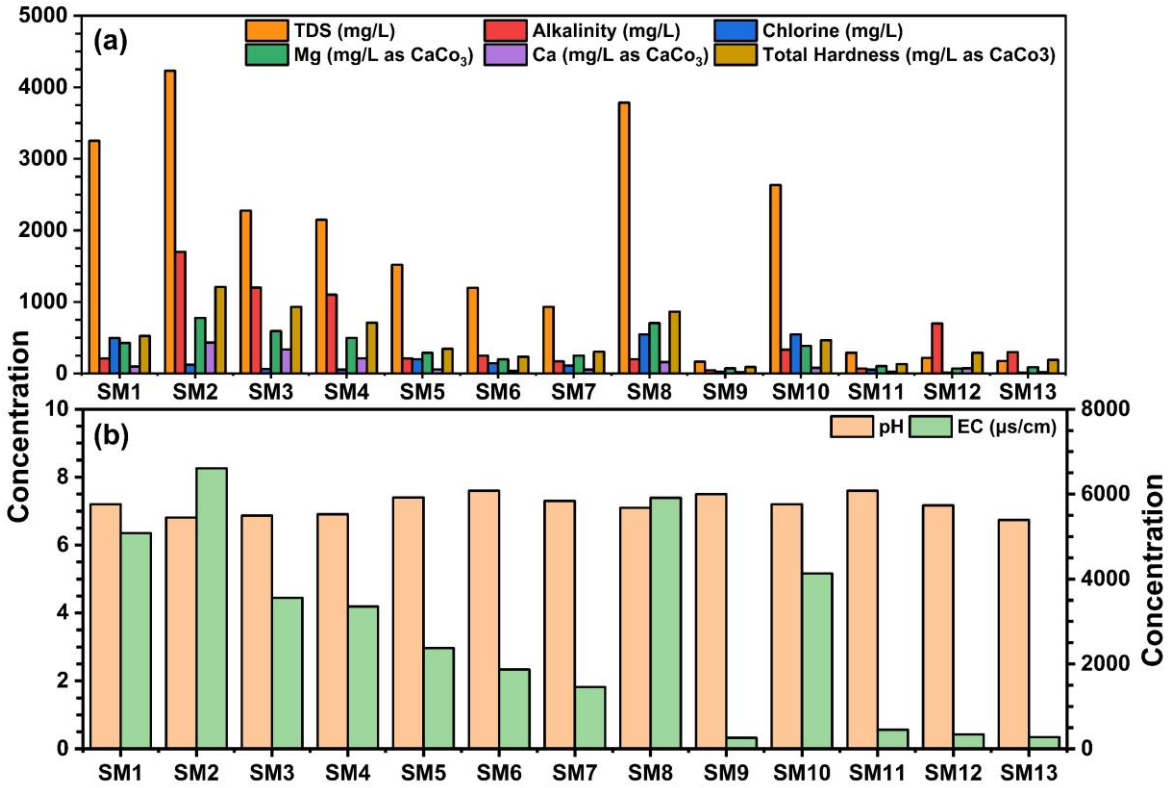


Figure 4.6 Physicochemical characteristics of submersible (SM1-SM13) groundwater samples of (a) EC, TDS, Major Ions, Hardness, and (b)pH.

The results of BH, handpump, and submersible pumps revealed that the water quality in the regions surrounding the landfill is loaded with higher concentrations of pollutants, including heavy metals, which are found in almost all the water samples collected for the study.

4.3.5 IDENTIFICATION OF CONTAMINATION IN SURFACE WATER SAMPLE

Surface water samples were collected from Bhalswa Lake, located near the Bhalswa dumpsite. The surface water quality analysis near the landfill region revealed that the concentration of Dissolved Oxygen (DO) is critically low in both the samples collected (1.56 and 2.42 mg/L) (Figure 4.6). The lower level of DO is indicative of a higher BOD level, which resulted from the untreated organic waste or leachate infiltration from the landfill sites. The DO value is approximately half the required permissible limit of 4 mg/L of CPCB standards. In both water samples, TDS levels exceeded the permissible limit (500 mg/L) (Figure 4.6). Table 4.4 illustrates the physicochemical properties of surface water samples.

The EC values ranged from 713 to 816 µs/cm due to higher dissolved ions, resulting in lower resistivity. The salinity levels (0.65 and 0.70) indicate that the conversion from freshwater to brackish water makes it unsuitable for domestic and drinking purposes. The parameters, such as magnesium, Total Hardness, and alkalinity, were found to exceed the IS10500:2012 permissible limits for surface water quality. The physicochemical profile of the Bhalswa lake revealed that its water quality is influenced by landfill leachate infiltration, resulting in eutrophic to hypertrophic conditions in the freshwater sources.

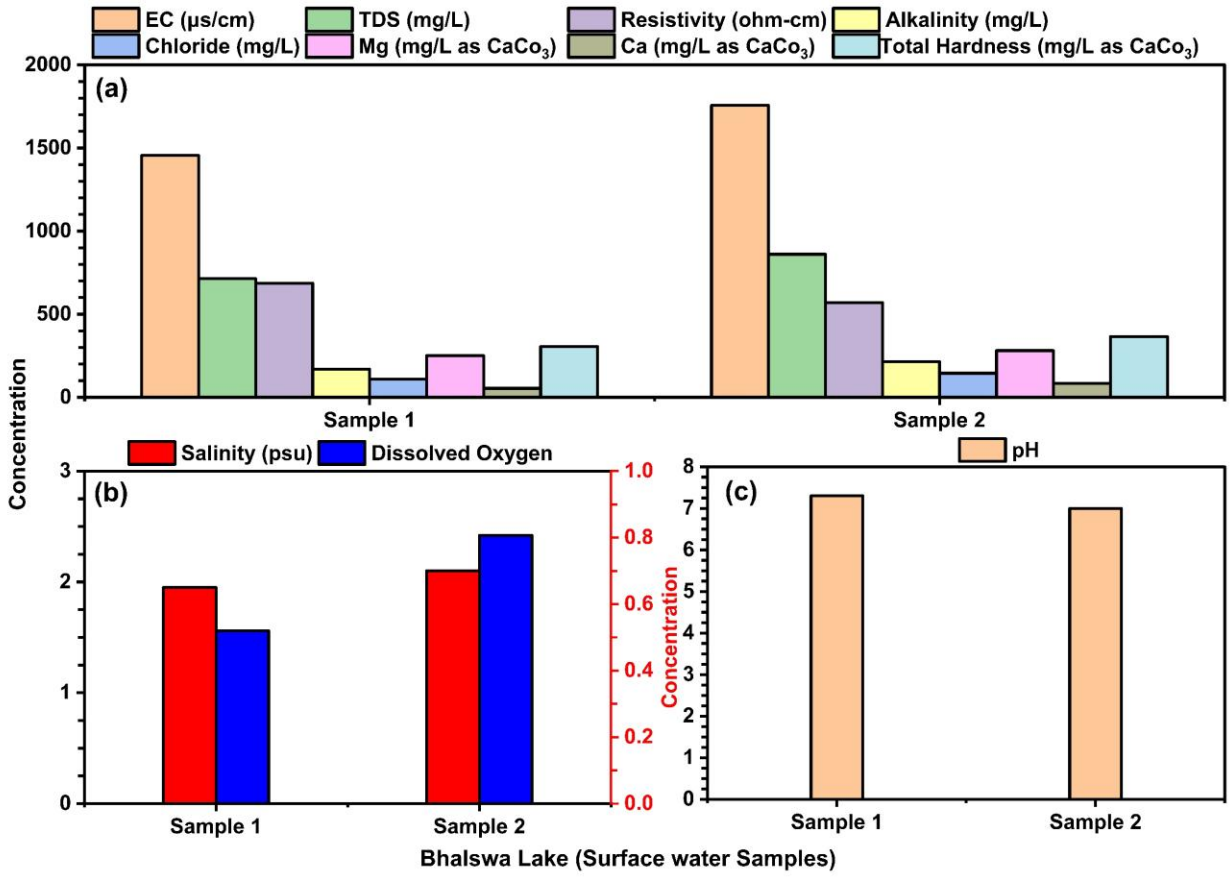


Figure 4.7 Physicochemical Characteristics of Surface Water Samples from Bhalswa Lake.

Table 4.5 Physicochemical analysis of surface water samples

Bhalswa Lake (Surface water Samples)	Sample 1	Sample 2
EC (µs/cm)	1455	1756
TDS (mg/L)	713.9	860.9
Resistivity (ohm-	686.9	569.5
Alkalinity (mg/L)	170	215
Chloride (mg/L)	110.121	145.1595
Magnesium (mg/L)	251	281
Calcium (mg/L)	54	84
Total Hardness	305	365
Salinity (psu)	0.65	0.7
pH	7.3	7
Dissolved Oxygen	1.56	2.42

4.4 WATER QUALITY INDEX (WQI) FOR ALL WATER SAMPLES

The weighted arithmetic water quality index (WA-WQI) is calculated for all 46 samples collected from the study area (Figure 4.7). The water quality index is a dimensionless scale that condenses complex physicochemical parameter data into a single value that represents overall water quality. The assessment of water quality (Balan et al., 2012) is determined by the values derived from the index, as illustrated in Table 4.5.

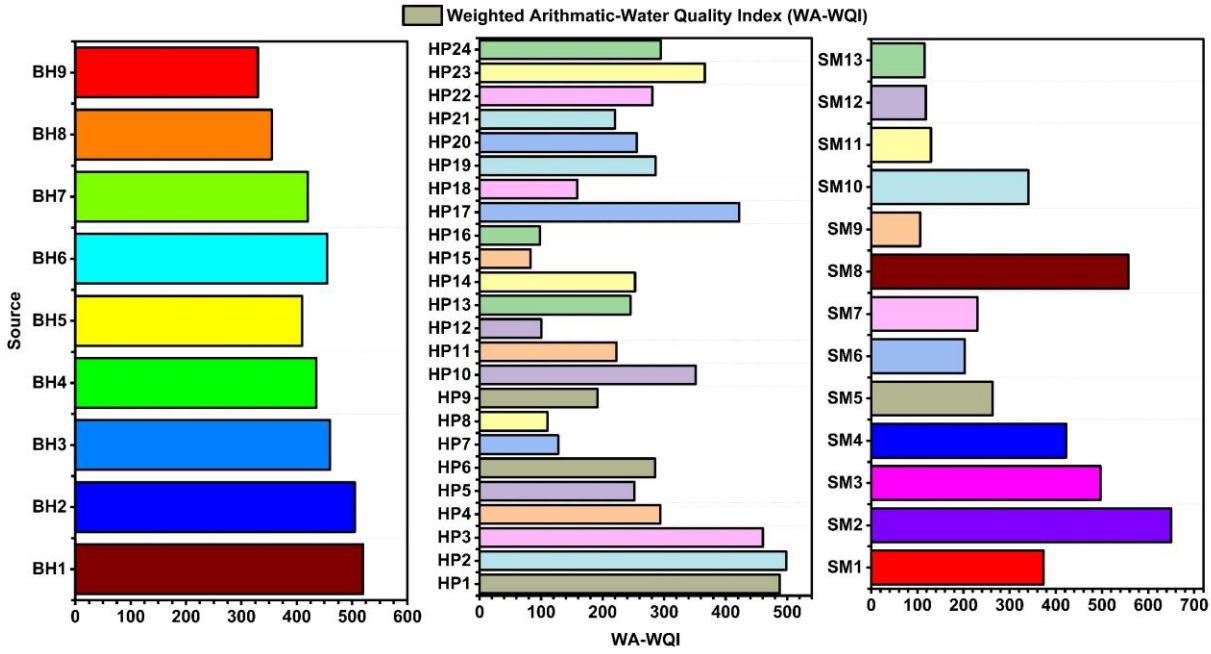


Figure 4.8 Weighted Arithmetic-Water Quality Index (WA-WQI) for borehole, hand pump, and submersible.

Table 4.6 Water quality indexing classification for all groundwater samples

Grading of Water quality	WQI value	Rating of water quality	Suitability to use	Sample
A	0-25	Excellent	Drinking, Industrial, Irrigation	0
B	25-50	Good	Drinking, Industrial, Irrigation	0
C	50-75	Poor	Industrial, Irrigation	0
D	75-100	Very Poor	Irrigation	3
E	>100	Unsuitable	Proper Treatment required	43

The results revealed that of the 46 collected samples, none were in the good category (excellent or good), and all were in the poor category. All the groundwater samples were in very poor condition (3) and unsuitable for use (43). The submersible water sources, identified as the most polluted among the other collected sources, with a maximum WQI value of 501.41, and a few water samples recorded in the poor category.

4.5 CONCLUSION

Groundwater samples from multiple sources around the Bhalswa landfill were analyzed to assess water quality in the region. The results revealed the severe degradation of the water quality of both ground and surface water. The concentration analysis of EC and TDS revealed that handpumps had the highest concentrations, followed by submersible pumps and boreholes. The TDS concentration in boreholes was found to be within the permissible limits. 5 times the permissible limit was found for Total hardness. Calcium and magnesium concentrations were also found to exceed the permissible limits. The total hardness concentration in submersible pumps was found to be the highest at 1210 mg/L, which is 5 times the permissible limit.

The Fe concentration was found to be 176.58-273.6 mg/L, and the Al concentration was also higher at 152.26 mg/L. Apart from groundwater, surface water sources were also found to be degraded, with higher salinity levels (0.65 to 0.70) and lower DO levels (1.56 and 2.42 mg/L). The heavy metal analysis for BH 1 to 9 revealed that, out of 12 parameters analyzed, only three were within safe limits; two were within the permissible range, and 7 exceeded both the permissible and acceptable range limits. The pollution levels in surface water indicate a transition from freshwater to brackish water. Magnesium, total hardness, and alkalinity analysis of surface water indicate eutrophic and hypertrophic conditions, making it unfit for use.

The water quality index further highlighted the severity of the degradation levels in the study area. The submersible water pump records the highest pollution concentration, with a WQI of 501.41,

followed by the handpump, with a WQI of 426.01. The study of the leachate sample indicated elevated electrical conductivity (EC) and total dissolved solids (TDS), along with heavy metal contamination, which was also detected in all water samples collected near the landfill site.

The groundwater quality assessment indicates that groundwater acts as the major environmental receptor in the landfill-affected area. The deterioration of physicochemical parameters and the presence of heavy metals in groundwater samples suggest the possible influence of leachate migration through the subsurface soil profile. The groundwater findings, when interpreted together with the leachate and soil contamination results, establish a clear source–pathway–receptor relationship. Bhalswa landfill leachate represents the contamination source, the soil profile acts as the migration pathway and intermediate sink, and groundwater functions as the final receptor affected by contaminant movement.

CHAPTER 5: DEVELOPMENT OF PREDICTIVE MODELLING FRAMEWORK FOR CONTAMINANT MIGRATION

5.1 INTRODUCTION

Groundwater and soil are essential natural resources that support human consumption, agricultural activities, ecological stability, and environmental sustainability. The quality of these resources is governed by their physical and chemical composition, which may change under the influence of natural processes and anthropogenic activities. Contamination of groundwater and soil has therefore become a significant environmental concern because deterioration in these resources can affect public health, plant growth, geotechnical stability, and long-term resource management. In regions affected by rapid urbanization, improper waste disposal, and unmanaged solid waste accumulation, contamination may occur through the movement of pollutants from surface sources into the subsurface environment (Bharadwaj et al., 2025 c).

Landfills and open dumps are important sources of environmental pollution because the decomposition of solid waste produces leachate, which contains dissolved organic matter, inorganic compounds, heavy metals, and other harmful constituents (Gaur et al., 2024). When leachate percolates through waste and migrates into surrounding soil and groundwater, the physicochemical characteristics of these resources may be altered. Soil can act as a sink for heavy metals and other contaminants, while groundwater can transport dissolved pollutants over wider subsurface zones. This interaction makes soil and groundwater contamination closely connected in environmental assessment, particularly where leachate movement, contaminant transport, and subsurface chemical variation are major concerns (Bharadwaj et al., 2025 a, c).

The assessment of groundwater quality commonly depends on hydro-chemical parameters such as pH, Electrical Conductivity, Chloride, Magnesium, Calcium, Potassium, and Total Hardness. These parameters are used to evaluate the suitability of water for drinking, irrigation, industrial

use, and other purposes (Vasanthavigar et al., 2010; Zahedi, 2017; Kumari and Rai, 2020; Maghrebi et al., 2021). The Water Quality Index provides a quantitative method for converting multiple water quality parameters into a single value that represents the overall condition of water. The Weighted Arithmetic Water Quality Index is one such approach that assigns relative importance to different physicochemical parameters and supports classification of water quality from excellent to unsuitable categories (Haghiabi et al., 2018; Adimalla et al., 2020; Lu and Ma, 2020; Nayan et al., 2020).

Soil contamination assessment involves the evaluation of physicochemical properties and heavy metal concentrations, including parameters such as pH, Electrical Conductivity, Total Dissolved Solids, chlorides, alkalinity, total hardness, magnesium, aluminium, manganese, iron, zinc, copper, cadmium, chromium, and lead (Alghamdi et al., 2021). Heavy metals are of particular concern because their mobility depends on chemical form, soil conditions, and contaminant pathways. Once accumulated in soil, these metals may affect soil behavior, plant uptake, and environmental quality (Kanmani et al., 2013). Therefore, understanding both horizontal and vertical variation in soil properties is important for identifying contamination patterns and assessing the influence of leachate migration.

Traditional laboratory-based assessment methods are useful but often time-consuming, costly, and resource-intensive when large-scale or continuous monitoring is required. Machine learning techniques provide an efficient alternative for predicting groundwater and soil contamination by identifying complex relationships among physicochemical parameters, heavy metals, and quality indicators (Bouzayani et al., 2014; Koda et al., 2020; Yoshida et al., 2002; Zairi et al., 2004). Models such as Artificial Neural Network, Random Forest, Support Vector Machine, XGBoost, Decision Tree, K-Nearest Neighbors, Logistic Regression, and hybrid approaches such as Random Forest with Genetic Algorithm have been used for environmental prediction tasks. These models

support regression and classification analysis, improve prediction accuracy, and assist in reducing dependence on repeated laboratory testing.

Machine learning also supports feature selection, optimization, and model validation, which are important for reliable environmental assessment. Statistical methods such as Pearson correlation can help identify influential parameters and reduce redundant inputs, while cross-validation, hyperparameter tuning, and model evaluation metrics improve the robustness of prediction models. In groundwater studies, machine learning can be used to estimate WQI or WAWQI and identify important hydro-chemical parameters. In soil studies, machine learning can be applied to predict heavy metal contamination and evaluate contaminant behavior under varying physicochemical conditions.

The combined application of machine learning in soil and groundwater studies provides a data-driven framework for environmental risk assessment, sustainable water quality monitoring, and contamination management. By integrating physicochemical analysis, heavy metal assessment, water quality indexing, feature optimization, and predictive modelling, machine learning approaches can enhance the efficiency and reliability of soil and groundwater contamination studies. This chapter therefore introduces the use of machine learning techniques for assessing and predicting contamination in soil and groundwater, with emphasis on leachate influence, physicochemical parameters, heavy metals, and water quality evaluation.

5.2 MATERIALS AND METHODS

5.2.1 STUDY AREA

The groundwater quality assessment was primarily centered around the Bhalswa landfill and nearby locations extending towards Delhi Technological University, Rohini, to assess the influence of landfill leachate on water quality. Additional groundwater records from Delhi and the

National Capital Region were incorporated to provide wider regional representation for WQI assessment and machine learning-based prediction. The broader region lies between 26.8042°N to 30.1380°N latitude and 75.5158°E to 78.2567°E longitude and includes Delhi and adjoining parts of Uttar Pradesh, Haryana, and Rajasthan (Figure 5.1). The geological setting comprises the Indo-Gangetic Alluvial Plains, the Aravalli Range, hard rock formations, and Quaternary Alluvium Plain, which influence groundwater occurrence and quality variation (CPCB 2024).

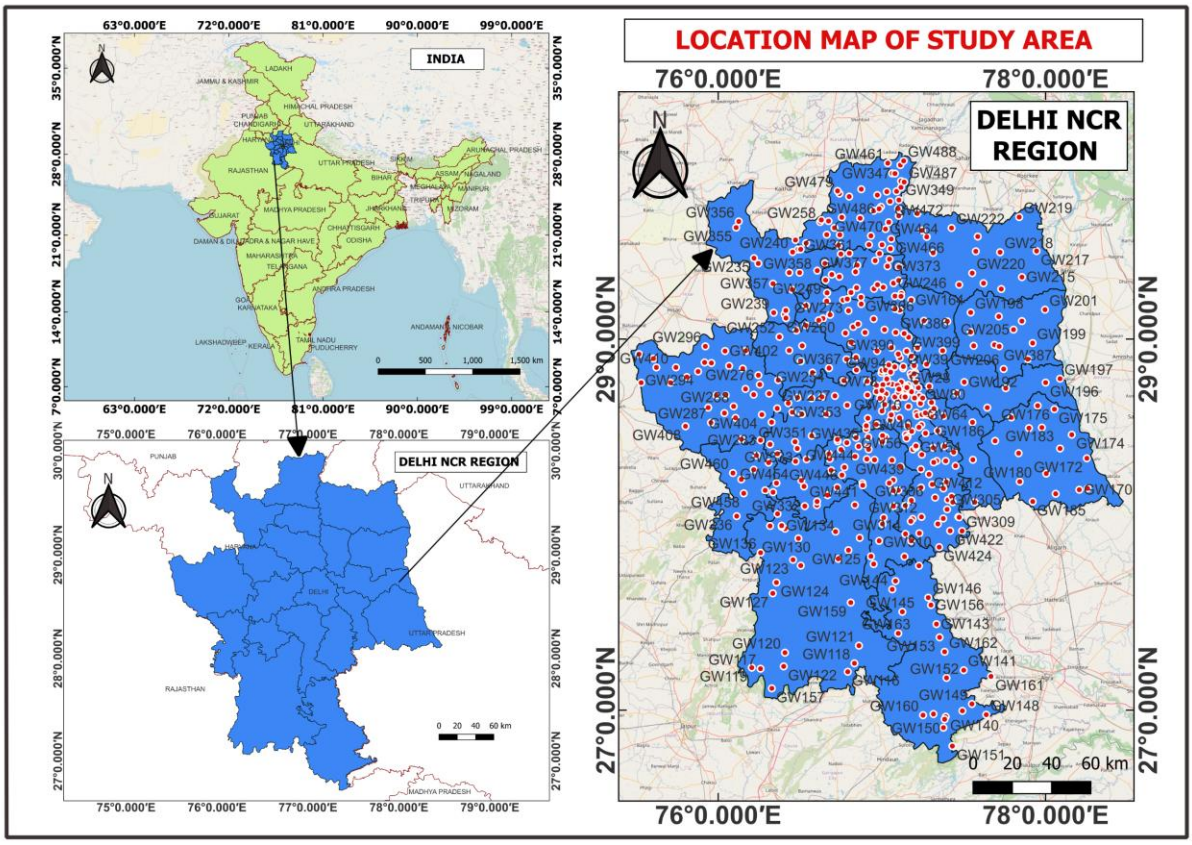


Figure 5.1 Ground water sample locations (Delhi, NCR (National Capital Region), India)

The soil contamination assessment was performed around the Bhalswa landfill in New Delhi, India. The landfill has been an active municipal solid waste disposal site since 1990 and has accumulated a large quantity of mixed and unsegregated waste over several decades. The site is located near Bhalswa Horseshoe Lake and lies approximately 10 km west of the Yamuna River. The landfill receives municipal solid waste from residential and industrial sources, and prolonged

waste accumulation has created conditions suitable for leachate generation and contaminant migration. The surrounding soil predominantly comprises sandy silt and clay loam, with alluvial origin due to location near the Yamuna floodplain. The soil study area (Figure 5.2) was selected to evaluate horizontal and vertical variation in physicochemical properties and heavy metal contamination associated with landfill leachate migration.

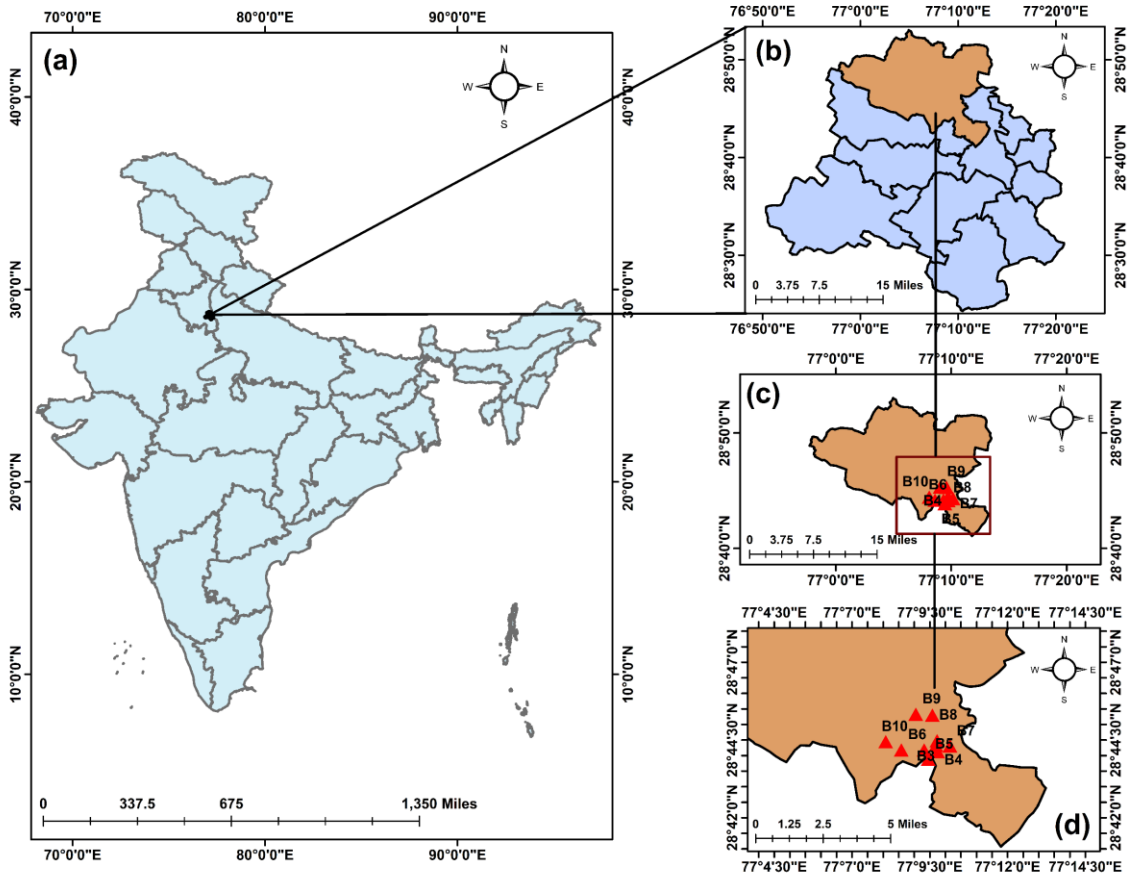


Figure 5.2 Study location Map (a) India (b) Delhi (c) North District Delhi (d) Bore Hole locations for soil sampling

5.2.2 GROUNDWATER SAMPLES

The groundwater quality assessment was based on 488 groundwater quality datasets. Out of these, 36 groundwater samples were collected from diverse locations around the Bhalswa landfill during November 2023 to January 2024 to assess the influence of landfill leachate on water quality. The

samples were collected through boreholes, hand pumps, and submersible pumps. The remaining 452 groundwater records were obtained from the Central Ground Water Board report for Delhi NCR. The combined groundwater dataset was used for calculating actual Weighted Arithmetic Water Quality Index values and for developing machine learning models for predicted WQI values (Bharadwaj et al, 2025 c,d).

Groundwater samples were collected in accordance with Indian standard codes ISO 5667-11:2009 and ISO 5667-3:2018. After collection, the samples were placed in an ice box and transported to the laboratory for further analysis. The samples were preserved at 4°C before laboratory testing. Seven hydro-chemical parameters were selected for groundwater quality assessment, namely pH, Electrical Conductivity, Chloride, Magnesium, Calcium, Potassium, and Total Hardness. The pH and Electrical Conductivity were measured using a multiparameter electrode. Chloride and total hardness were assessed by the titrimetric method as per APHA-EDTA. Magnesium, Calcium, and Potassium ions were measured using a flame photometer.

The results of the 488 groundwater datasets were used for model development and water quality assessment. Statistical summaries were prepared for the selected hydro-chemical parameters, and Pearson correlation coefficient analysis was performed to examine the influence of these parameters on WQI. The same dataset structure also supported feature reduction scenarios, where different combinations of input parameters were evaluated for accurate and cost-effective WQI prediction. The workflow for groundwater quality assessment using ML is illustrated in Figure 5.3.

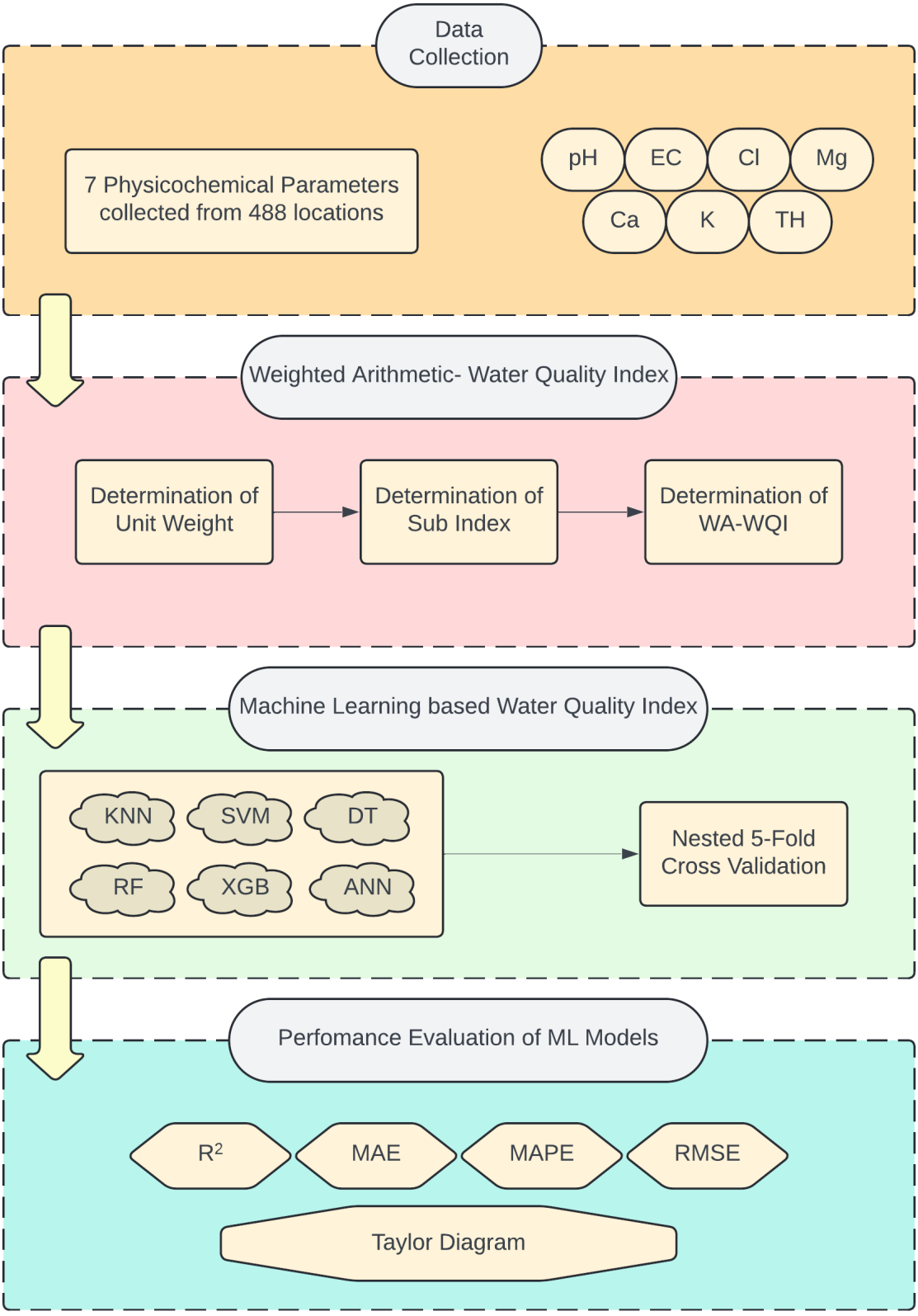


Figure 5.3 Flowchart for machine learning-based prediction of water quality

5.2.3 SOIL SAMPLING

Soil sampling was conducted around the Bhalswa landfill to assess the influence of leachate contamination on subsurface soil quality as described in C of oven driedhapter 3. The boreholes were drilled using percussion drilling in accordance with IS:1892. The soil samples were collected through borehole drilling and stored in airtight polythene zip-lock bags. After collection, the samples were transported to the laboratory for physicochemical and heavy metal analysis. The collected samples were oven-dried at 80°C for 24 hours and sieved through a 2 mm mesh to remove grass, pebbles, and other impurities.

For physicochemical analysis, 15 g of oven dried soil sample was mixed with 150 ml of distilled water in a conical flask at a 1:10 ratio. The flask was placed on an orbital shaker at 120 rpm for 5 hours. After shaking, the suspension was left undisturbed to allow soil settlement, followed by filtration using 25-micron filter paper. For highly turbid suspensions, centrifugation was performed at 4500 rpm for 5 minutes before filtration.

For heavy metal extraction, dried soil samples were heated for 12 to 24 hours at 100°C to 115°C. Approximately 1 to 2 g of dried soil was weighed and transferred to a 50 ml beaker. Concentrated HNO_3 was added, followed by heating on a hot plate at 100°C for 10 to 15 minutes. After cooling, additional HNO_3 was added and heating was continued until completion of reaction. Distilled water and H_2O_2 were added, and the mixture was reheated with incremental H_2O_2 addition until no further reaction was observed. The solution was then heated until the volume reduced to approximately 5 ml. After cooling, HCl was added and heating was continued for another 15 minutes. The fully digested solution was filtered and diluted to a final volume of 100 ml using distilled water in a volumetric flask.

The prepared soil solution was used for heavy metal analysis through Microwave Plasma-Atomic Emission Spectroscopy. A multiparameter electrode was used to measure pH, Electrical

Conductivity, resistivity, and salinity. Titration methods were used for alkalinity, chloride content, and total hardness, including calcium and magnesium. The soil contamination assessment included physicochemical parameters and heavy metals such as magnesium, aluminium, manganese, iron, zinc, copper, cadmium, chromium, and lead.

5.2.4 WEIGHTED ARITHMETIC WATER QUALITY INDEX

The Weighted Arithmetic Water Quality Index was used to assess overall groundwater quality by combining multiple physicochemical parameters into a single index value. The selected parameters for WAWQI calculation were pH, Electrical Conductivity, Chloride, Magnesium, Calcium, Potassium, and Total Hardness. The method assigns unit weights to parameters according to their relative significance in water quality assessment and categorizes water quality based on calculated WQI values.

The WAWQI calculation followed four steps as described in Chapter 4 from Equation 4.3 – 4.6. First, the quality rating scale for each parameter was calculated using the actual value of the water quality parameter, the ideal value, and the recommended standard value. The ideal value was considered as 7 for pH and zero for all other parameters. Second, the proportionality constant was calculated. Third, the unit weight for each parameter was determined. Fourth, the Water Quality Index was calculated through weighted aggregation of the quality rating scale and unit weight. The calculated WQI values were used as target values for machine learning-based prediction. Figure 5.4 shows the flowchart for developing WA-WQI.

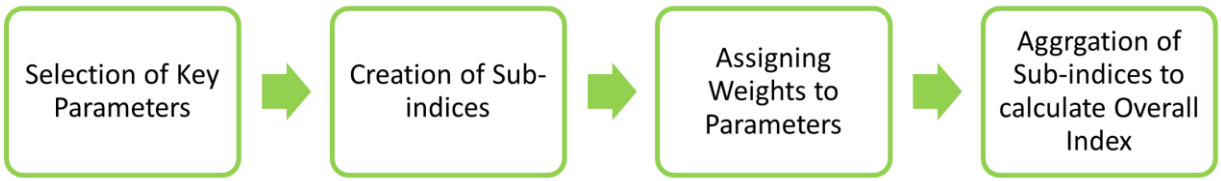


Figure 5.4 Flowchart of steps for developing WA-WQI

5.2.5 MACHINE LEARNING MODELS

Machine learning models were applied for groundwater quality prediction and soil heavy metal contamination assessment. For groundwater analysis, six supervised regression models were used to predict WA-WQI values: Decision Tree, Artificial Neural Network, Random Forest Regressor, Support Vector Machine, XGBoost, and K-Nearest Neighbors. The objective was to estimate continuous WQI values from selected hydro-chemical input parameters.

For soil contamination analysis, machine learning models were applied to predict heavy metal contamination associated with landfill leachate migration. The soil manuscript compared Random Forest, XGBoost, Support Vector Machine, Artificial Neural Network, and Logistic Regression with a hybrid Random Forest-Genetic Algorithm model (RF-GA). The hybrid RF-GA model was developed for feature selection, optimization, and improved prediction performance.

5.2.5.1 DECISION TREE

A Decision Tree is a non-parametric model that does not presume any specific distribution of the data, making it capable of capturing complex, non-linear interactions among water quality parameters without explicitly defining mathematical functions. The process involves recursively splitting the dataset at each node into two child nodes according to feature values, with the objective of minimizing the Mean Squared Error (MSE) of the target variable. The procedure persists until a stopping requirement such as maximum tree depth or minimum sample size per node is satisfied. For novel predictions, the model allocates the mean target value of the leaf node corresponding to the input data. Data scaling and normalization are not essential for DT, as pH is unitless and parameters such as EC (µS/cm), Cl^-, Mg^{2+}, Ca^{2+}, K^+, and TH are measured in mg/L. Decision Trees can handle both numerical and categorical data types and create a flowchart-like

structure, making them useful as a fundamental component of the predictive modelling framework to predict groundwater quality. (Abba et al., 2024).

5.2.5.2 ARTIFICIAL NEURAL NETWORK

ANN is a machine learning model inspired by the information processing mechanisms of biological neural networks in the human brain. The Multilayer Perceptron is a commonly utilized ANN architecture for classification and regression tasks. It is made up of multiple layers an input layer, one or more hidden layers, and an output layer with each layer containing perceptive units known as neurons. ANN performed better than other models when handling large numbers of datasets for groundwater quality forecasting (Nordin et al., 2021). Due to its ability to model complex non-linear relationships. ANN identifies hidden relationships in historical data that are beyond the capabilities of human cognitive analysis, and can demonstrate any continuous function with a sufficient number of neurons, regardless of the complexity of the WQI-parameter relationship. (Abd-Elmaboud et al., 2024). In soil contamination assessment, ANN was used as a comparison model for predicting heavy metal contamination.

5.2.5.3 RANDOM FOREST

Ensemble methods enhance the predictive accuracy of an individual model by training several models and merging their predictions. Random Forest is one such ensemble method that creates multiple decision trees and predicts their aggregate (Arafa et al., 2024), Thereby reducing overfitting and handling high-dimensional data. RF combines multiple decision trees as base models by integrating bagging technique and random feature selection. Bagging (bootstrap aggregating) randomly resamples the original data to create training datasets with replacements, known as bootstrap samples. Each tree in RF is trained on a distinct bootstrap sample, using a

randomly selected subset of the predictor (feature) variables. RF is particularly successful at dealing with imbalanced datasets, as polluted samples are more common than clean water samples. The RF model was optimized using parameters including the maximum depth of each decision tree, the number of decision trees, and the minimum number of samples required to split an internal node and a leaf node (Sathiyamurthi et al., 2024). Random Forest was used for both groundwater quality prediction and soil contamination assessment. Random Forest also served as the base model for the RF-GA hybrid approach.

5.2.5.4 SUPPORT VECTOR MACHINE

SVM, introduced to the computer science field by Vladimir Vapnik in 1995, is a widely used and highly effective supervised machine learning algorithm. It is a robust technique that uses kernel functions to transform data into a higher-dimensional space, facilitating linear separation (El-Said Salem et al.,2025). It provides high accuracy in delivering complex relationships between topography, hydrology, and geology for groundwater mapping. Numerous studies demonstrate that SVM performs better than traditional techniques for water quality predictions even with constrained training data (Jamshidzadeh et al.,2024). The SVM model in this study is configured with the following parameter C = 0.1, gamma = "scale", and kernel = "linear". These parameters define the margin width of the SVM and the selection of the kernel, which enhances the accuracy of WQI predictions. SVM was applied for groundwater WQI prediction because the model can capture non-linear associations among physicochemical parameters and WQI values. The model was also used as a comparison algorithm in the soil contamination study.

5.2.5.5 XGBOOST

XGBoost became popular in ML competitions after being developed across multiple languages such as Python, R, and Java. It is essentially an advanced ensemble of decision trees with a modified gradient boosting framework. Boosting combines several weak models sequentially to create a strong model, with each new model focusing on correcting the errors of the previous ones. Every time a tree is added, it learns a new function, and XGBoost is particularly effective for early warning systems as it accurately forecasts future groundwater quality. XGBoost is versatile and effective in water resources engineering (Niazkar et al., 2024). XGBoost was applied for groundwater WQI prediction and soil contamination assessment because of strong performance in complex environmental datasets.

5.2.5.6 K-NEAREST NEIGHBORS

First developed by Evelyn Fix and Joseph Hodges in 1951 as part of military research, KNN is a simple supervised machine learning algorithm that can be applied for classification or regression tasks. KNN is a lazy and non-parametric algorithm, meaning it does not need a pre-built model for new data points and instead stores all the data at hand, organizing additional information based on similarity. KNN utilizes local neighborhood knowledge for predictions, determining the quality of a water sample based on the quality of its nearest neighbors. This yields reliable results that are also advantageous for spatial water quality modelling (Srisuradetchai et al., 2024).

5.2.5.7 LOGISTIC REGRESSION

Logistic Regression was used in the soil contamination study as a comparison model for heavy metal contamination prediction. The model estimates class membership through a probabilistic relationship between input variables and target classes. Logistic Regression was compared with

Random Forest, XGBoost, Support Vector Machine, ANN, and RF-GA to evaluate the predictive advantage of the hybrid optimization model.

5.2.5.8 RANDOM FOREST-GENETIC ALGORITHM MODEL (RF-GA)

The Random Forest-Genetic Algorithm model was developed for predicting heavy metal contamination in soil. Random Forest was used as the base predictive model, while Genetic Algorithm was used for optimization and feature selection. Genetic Algorithm is based on the principles of natural selection and evolutionary theory. The process begins with an initial population of possible solutions represented as chromosomes. Each chromosome contains genes corresponding to candidate feature combinations or model parameters.

The Genetic Algorithm follows selection, crossover, mutation, elitism, and termination stages. Selection identifies superior chromosomes for generating the next population. Crossover combines genes from parent chromosomes to produce offspring. Mutation introduces variability by modifying selected genes and helps reduce premature convergence. Elitism preserves superior chromosomes in the next generation. The termination condition may be based on maximum number of generations, satisfactory fitness value, or convergence behavior. The optimized feature set obtained through Genetic Algorithm was integrated with Random Forest to enhance prediction accuracy and reduce redundancy in soil contamination assessment.

5.2.6 RESAMPLING TECHNIQUES AND HYPERPARAMETER OPTIMIZATION

Resampling and hyperparameter optimization were used to improve model reliability and reduce dependence on a single train-test split. For groundwater quality prediction, nested 5-fold cross-validation and grid search were applied to optimize machine learning models. The groundwater

dataset was divided into five subsets. In each iteration, four subsets were used for model training and one subset was used for model evaluation. The process was repeated until each subset served as the evaluation fold. Average performance metrics were used for model assessment.

Grid search was used to identify suitable hyperparameter combinations for each model. Cross-validation improved data utilization and supported robust evaluation of model performance. For soil contamination prediction, multiple training ratios were used to assess learning stability and generalization. The RF-GA model used Genetic Algorithm-based optimization to enhance feature selection and model performance.

5.2.7 DATA ANALYSIS AND FEATURE SELECTION

Data analysis was performed through statistical summaries, Pearson correlation analysis, pair plots, and feature selection methods. For groundwater quality prediction, statistical summaries were prepared for hydro-chemical parameters and WQI values. Pearson correlation coefficient was used to examine the strength and direction of linear relationships between physicochemical parameters and WQI. Pearson correlation coefficient, denoted as r, is a widely used statistical metric that assesses the strength and direction of the linear relationship between two continuous variables. The correlation strength increases with the absolute value of r, which is confined to the interval [−1, 1]. When r > 0, the variables exhibit a positive correlation; when r < 0, a negative correlation is observed; and if r = 0, there exists no linear correlation among the variables. The formula is given in equation 5.1.

$$r = \frac{\Sigma(X_i - X)(Y_i - \bar{Y})}{\sqrt{[\Sigma(X_i - X)^2 \times \Sigma(Y_i - \bar{Y})^2]}} \qquad \textbf{(5.1)}$$

Where,

- X_i and Y_i are the values of X and Y for the i-th individual

- $\bar{X}$ and $\bar{Y}$ are the respective means
- N is the number of data points.

Feature selection was performed to identify influential input parameters for WQI prediction. Pearson correlation coefficient was used because the method is fast, simple, independent from the modelling framework, and suitable for datasets containing multiple predictor variables. Parameter reduction scenarios were created to evaluate whether fewer input features could maintain reliable WQI prediction. The groundwater study evaluated all-feature and reduced-feature cases using combinations of pH, Electrical Conductivity, Magnesium, Potassium, Total Hardness, Calcium, and Chloride.

For soil contamination assessment, Pearson correlation matrix heatmaps were used to examine interrelationships among independent variables and target heavy metals such as Cadmium, Chromium, and Lead. Variables including pH, Electrical Conductivity, Total Dissolved Solids, sodium, potassium, calcium, magnesium, chloride, nitrate, sulphate, fluoride, carbonate, bicarbonate, alkalinity, iron, zinc, copper, aluminium, manganese, depth, and distance from the contamination source were examined. Chi-square-based feature selection was used with correlation insights to identify meaningful predictors and reduce redundancy. This procedure supported improved RF-GA model performance in soil heavy metal contamination prediction.

5.2.8 MODEL EVALUATION METRICS

Model evaluation metrics were used to compare predicted values with actual values and to assess model performance. For groundwater WQI prediction, coefficient of determination, Mean Absolute Error, Mean Absolute Percentage Error, Root Mean Square Error, and Centered Root Mean Square Error were used. The coefficient of determination indicates the proportion of

variance explained by the model. Mean Absolute Error measures the magnitude of prediction errors without considering direction. Mean Absolute Percentage Error expresses prediction error as a percentage of actual values. Root Mean Square Error measures the average magnitude of squared errors and is expressed in the same unit as the target variable. Centered Root Mean Square Error (Equation 5.6) was used for Taylor diagram assessment.

Coefficient of Determination (R^2)

R^2 (Equation 5.2) is the ratio of the explained variation compared to the total variation. It is a dimensionless quantity. Values closer to 1 suggest greater model performance.

$$\boldsymbol{R^2 = 1 - \frac{[\Sigma(y_i - \hat{y}_i)^2]}{[\Sigma(y_i - \bar{y})^2]}} \qquad \textbf{(5.2)}$$

Where, y_i are actual observed values, $\hat{y}_i$ are predicted values, and $\bar{y}$ is the mean of observed values. (Worst value = $-\infty$; best value = +1)

Mean Absolute Error (MAE)

MAE (Equation 5.3) measures the magnitude of inaccuracies in a set of forecasts regardless of their direction, expressed in the same unit as the target variable.

$$\textbf{MAE} = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \qquad \textbf{(5.3)}$$

where N is the number of observations. (Worst value = ∞; best value = 0)

Mean Absolute Percentage Error (MAPE)

MAPE (Equation 5.4) is a popular measure of forecast accuracy due to its scale independence and clarity. It calculates the error as a percentage of the actual value.

$$\mathbf{MAPE} = \frac{\mathbf{1}}{\mathbf{N}} \sum_{\mathbf{i=1}}^{\mathbf{N}} \left| \frac{\mathbf{y_i - \hat{y}_i}}{\mathbf{y_i}} \right| \times \mathbf{100} \qquad \textbf{(5.4)}$$

(Worst value = ∞; best value = 0)

Root Mean Square Error (RMSE)

RMSE (Equation 5.5) is a measure of average error based on the sum of squared errors. It is expressed in the same unit as the target variable.

$$\mathbf{RMSE} = \sqrt{\frac{\mathbf{1}}{\mathbf{N}} \sum_{\mathbf{i=1}}^{\mathbf{N}} (\mathbf{y_i - \hat{y}_i})^{\mathbf{2}}} \qquad \textbf{(5.5)}$$

(Worst value = ∞; best value = 0).

$$\mathbf{Centered\ RMSE} = \sqrt{\frac{\mathbf{1}}{\mathbf{N}} \sum_{\mathbf{i=1}}^{\mathbf{N}} [(\mathbf{a_i - \bar{a}}) - (\mathbf{b_i - \bar{b}})]^{\mathbf{2}}} \qquad \textbf{(5.6)}$$

- a_i is the i-th actual value
- b_i is the i-th predicted value
- $\bar{a}$ is the mean of the actual values
- $\bar{b}$ is the mean of the predicted values
- N is the number of observations

For soil contamination prediction, accuracy, precision, recall, F1-score, sensitivity, specificity, False Negative Rate, and False Positive Rate were used to assess classification performance. Accuracy measures the proportion of correct predictions. Precision evaluates the proportion of positive predictions that are correct. Recall and sensitivity measure the proportion of actual positive cases correctly identified. Specificity measures the proportion of actual negative cases correctly identified. False Negative Rate and False Positive Rate evaluate misclassification errors.

These metrics were used to compare RF-GA with Random Forest, XGBoost, Support Vector Machine, ANN, and Logistic Regression.

5.2.9 TAYLOR DIAGRAM AND CONVERGENCE ANALYSIS

Taylor diagrams were used to assess the accuracy of machine learning models in forecasting WQI. A Taylor diagram is a polar graph that represents model performance relative to observed values using standard deviation, Pearson correlation coefficient, and centered root mean squared error. The distance between a model point and the reference point indicates model performance. A model closer to the reference point has standard deviation similar to observed values, higher correlation, and lower centered root mean squared error. Taylor diagram analysis was used to rank groundwater prediction models and identify the best-performing model for validation.

Convergence analysis was used to evaluate optimization performance in the soil contamination study. The analysis examined reduction in cost function values across iterations. The cost function quantifies the difference between predicted and actual target values, and lower cost values indicate improved model performance. The RF-GA model was compared with Random Forest, XGBoost, Support Vector Machine, ANN, and Logistic Regression. The RF-GA model demonstrated rapid cost reduction and lower final cost function values, indicating enhanced convergence and improved optimization capability for heavy metal contamination prediction in soil.

5.3 RESULTS AND DISCUSSION

5.3.1 ASSESSMENT OF GROUNDWATER QUALITY USING PHYSICOCHEMICAL PARAMETERS

The groundwater quality assessment was performed using seven hydro-chemical parameters, namely pH, Electrical Conductivity, Chloride, Magnesium, Calcium, Potassium, and Total Hardness. The Weighted Arithmetic Water Quality Index (WA-WQI) was calculated using Equations (4.3) to (4.6). The WA-WQI values range from 21.08 to 1630.401, indicating a broad spectrum of water quality in the study region, from excellent to unsuitable category. The Water Quality Index (WQI) is an essential instrument for evaluating water appropriateness for diverse purposes such as consumption, irrigation, and industrial usage. The grading system indicates that the majority of samples are classified as unsuitable to poor quality categories, reflecting the significant contamination of water in the investigated region. The majority of groundwater samples from the study region are unsuitable for drinking purpose. The 488 samples are categorized into grades A through E, indicating the potential for groundwater utilization based on water quality ratings, as detailed in Table 5.1.

Based on the analysis of 488 groundwater samples, the WQI classification reveals significant water quality concerns. Only one sample falls under the "Excellent" category (WQI 0–25), making it safe for all uses. Thirteen samples are rated "Good" (WQI 25–50), suitable for drinking, irrigation, and industrial purposes. A significant number of samples (130) are classified as "Poor" (WQI 50–75), signifying that the water is unsuitable for consumption but may still be utilized for irrigation and industrial purposes. The highest number of samples, 224, belong to the "Very Poor" category (WQI 75–100), restricting their application solely to irrigation. Alarmingly, 120 samples exceed a WQI of 100, making them unsuitable for any use without proper treatment. These findings highlight the urgent need for groundwater monitoring and treatment strategies to improve water quality and ensure safe usage. The statistical summary of the gathered physicochemical

parameters and the derived water quality index is presented in Table 5.2, containing the minimum, maximum, mean, total sample count, 25th percentile, 75th percentile, median (50th percentile), and standard deviation of the water quality parameters. The parameters Hardness (488.895 ± 553.452 mg/L), Cl (407.261 ± 686.812 mg/L), Mg (98.411 ± 112.156 mg/L), and K (20.340 ± 50.745 mg/L) exceed the BIS 10500:2012 values for drinking water. However, EC (2307.563 ± 2716.073 µS/cm) and Ca (62.420 ± 87.886 mg/L) are below the standard limit.

Table 5.1 Class-wise Distribution of WQI

No. of Samples	WQI Value	Rating of Water Quality	Usage Possibility	Grading
1	0–25	Excellent	Drinking, Industrial, Irrigation	A
13	25–50	Good	Irrigation, Industrial, Drinking	B
130	50–75	Poor	Irrigation, Industrial	C
224	75–100	Very Poor	Irrigation	D
120	>100	Unsuitable	Proper Treatment Required	E

Table 5.2 Statistical Summary of Geochemical Parameters from 488 Groundwater Samples after Water Quality Index Assessment

Parameter	Count	Mean	Min	Max	SD	25th Percentile	Median (50th)	75th Percentile	Standard Value
pH	488	8.28	6.90	9.24	0.47	7.96	8.34	8.68	6.5–8.5 (BIS, 2012)
EC (µS/cm)	488	2307.56	39.08	24060	2716.07	775.25	1397.50	2873.50	2500 (EU, 2023)
Cl^- (mg/L)	488	407.26	7.10	5388	686.81	50	145.58	483	250 (BIS, 2012)
Mg^{2+} (mg/L)	488	98.41	2	846	122.16	29	49	109.75	30 (BIS, 2012)
Ca^{2+} (mg/L)	488	62.42	4.40	801	87.89	18	32	68	75 (BIS, 2012)
K^+ (mg/L)	488	20.34	0	520	50.75	3	6	14.07	12 (WHO, 2011)
TH (mg/L)	488	488.90	33	3814	553.45	186	290	534.50	200 (BIS, 2012)
WQI	488	148.24	21.09	1630.40	163.55	71.14	92.85	156.20	—

The average pH value in the area is 8.277 ± 0.467, suggesting that the majority of the groundwater is alkaline. Total hardness refers to the concentration of calcium and magnesium ions, which can influence the taste of water and lead to scale buildup. High levels of chloride concentration may indicate pollution or increased salinity. While magnesium is an essential mineral, excessive amounts can impact taste and health. Potassium is generally beneficial in small quantities, but high levels can be detrimental. The average and standard deviation of the WA-WQI for the basin are 148.241 ± 163.555. All the calculated values have been compared with different standards such as the BIS-10500 (2012) code, European Union (Drinking Water) Regulations 2023, and World Health Organization (2011) Guidelines for Drinking-Water Quality.

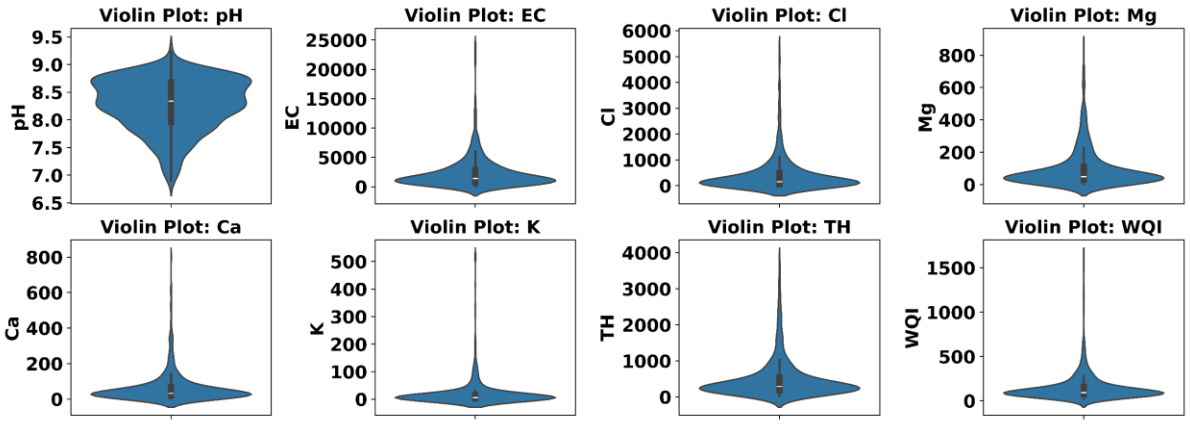


Figure 5.5 Distribution of geochemical water quality parameters used for machine learning models from 488 samples in Delhi and its surrounding region, India

Figure 5.5 illustrates the variability of geochemical parameters used to calculate the water quality index in the study region. The pH range falls between 7 and 9, as shown in the violin plot, indicating that the water is either neutral or alkaline in nature. This study shows higher electrical conductivity, which correlates with increased contaminants or dissolved salts in the water samples. The majority of data indicates poor electrical conductivity values. Excessive chloride concentration produces a salty taste; hence, elevated Cl^- levels are harmful for human health.

Magnesium is the primary contributor to water hardness. The distribution of Magnesium in the violin plot is widely dispersed, and it is typically concentrated at lower values. Calcium is also significant for water hardness, and excessive calcium might lead to scaling issues; here, the majority of values are low, but several outliers indicate elevated calcium levels. Potassium levels are predominantly in the low range. Hardness exhibits low levels in most locations, with outliers in a few areas.

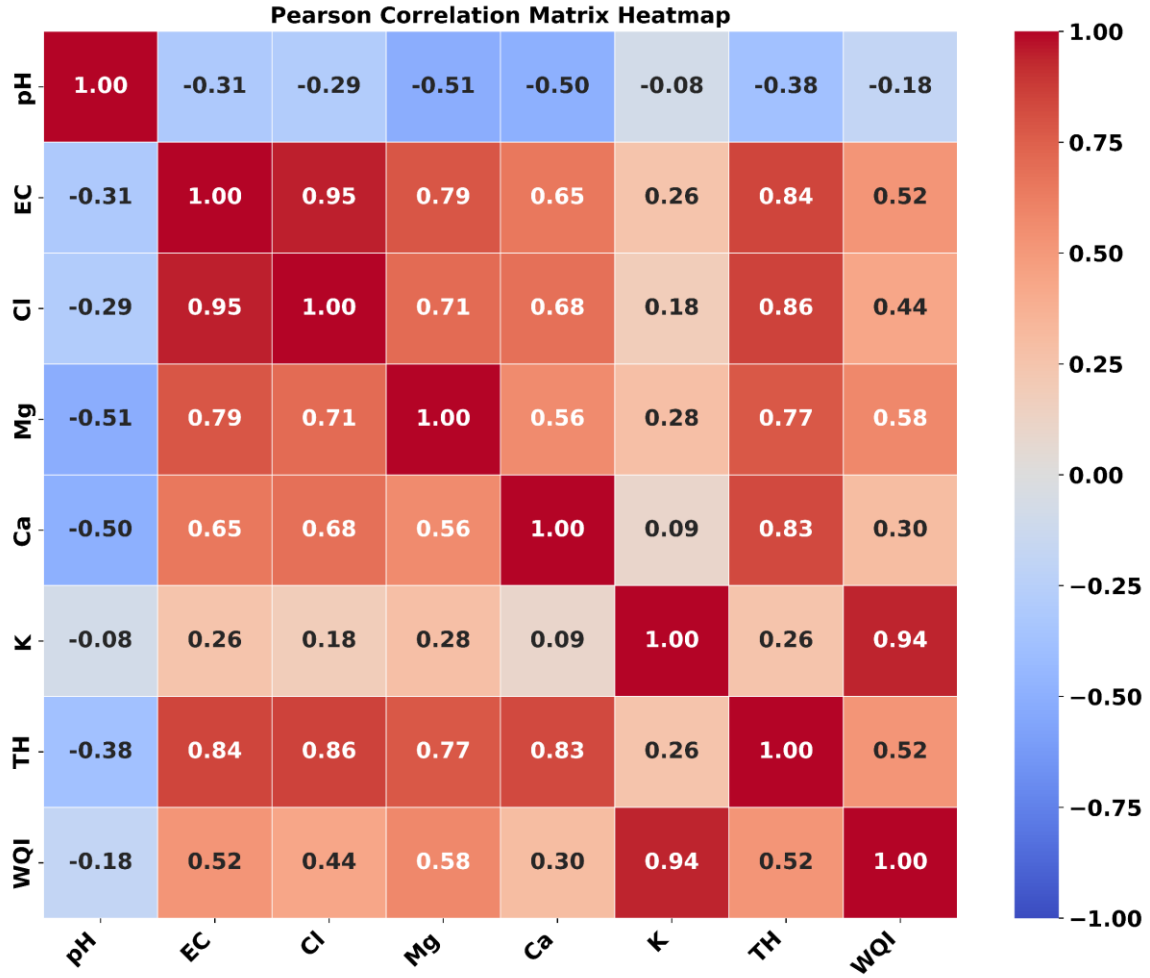


Figure 5.6 Pearson Correlation Matrix Heatmap for all 7 features and WQI

The Pearson Correlation Matrix (Figure 5.6) was used to examine relationships between physicochemical features and WQI. The range of the correlation coefficient (r) is from −1 to 1. The relationship is directly proportional when the value is positive and inversely proportional

when the value is negative; a higher coefficient value indicates a stronger relation (Ling et al., 2024). The top six features strongly correlated with WQI are K (r = 0.94), Mg (r = 0.58), TH (r = 0.52), EC (r = 0.52), Cl (r = 0.44), and Ca (r = 0.30). A negative correlation with WQI is observed for pH (r = −0.18). These findings indicate that Potassium (K) plays the most significant role in determining water quality in the study region, whereas pH shows a weak relation. This correlation hierarchy also formed the basis for the feature reduction scenarios explored in the subsequent analysis (Bharadwaj et al., 2025 c).

5.3.2 GEOSPATIAL ANALYSIS OF PHYSICO-CHEMICAL ATTRIBUTES AND EVALUATED WQI

Inverse Distance Weighting represents the spatial distribution of WQI across a region, using an interpolation method to estimate water quality between sampling points (Figure 5.7). The pH map indicates a zonal variance in the study area. The Centre and specific south-western regions exhibit low pH levels (red-yellow color), signifying that the water is quite acidic. The northern and boundary regions exhibit elevated pH levels (green shades), signifying an alkaline nature. Central regions and some southeastern regions exhibit elevated electrical conductivity (orange-red color), indicating a higher concentration of dissolved particles in the water. The north and west regions include comparatively lower salinity water (green zones). The chloride spatial map indicates elevated chloride concentrations in the central, south-central, and certain eastern regions (orange-red), which may signify contamination or human activity. Chloride levels are generally lower on the north side (green). Magnesium levels are high in the middle and south-west regions, while the rest have moderate to low levels. Calcium exhibits a higher concentration in the Centre and south-west (orange-red shades). It is deficient in the northeastern and certain western boundaries (green zones). The potassium map indicates elevated values in the central and southeastern regions (orange-yellow), whereas other places exhibit low potassium levels (green), common in

groundwater. The pattern of total hardness is consistent, with increased hardness (red-orange color) observed in the central and south-western regions, indicating hard water. Other places exhibit mild to moderate water quality (green to yellow).

The WQI map (Figure 5.7) shows the overall water quality in the area. Central and south-western regions are problematic zones where the Water Quality Index (WQI) is elevated (red color), indicating poor water quality. Northern and specific eastern locations have superior quality (green color) and low WQI. The different color gradient indicates WQI categories based on the quality of water index.

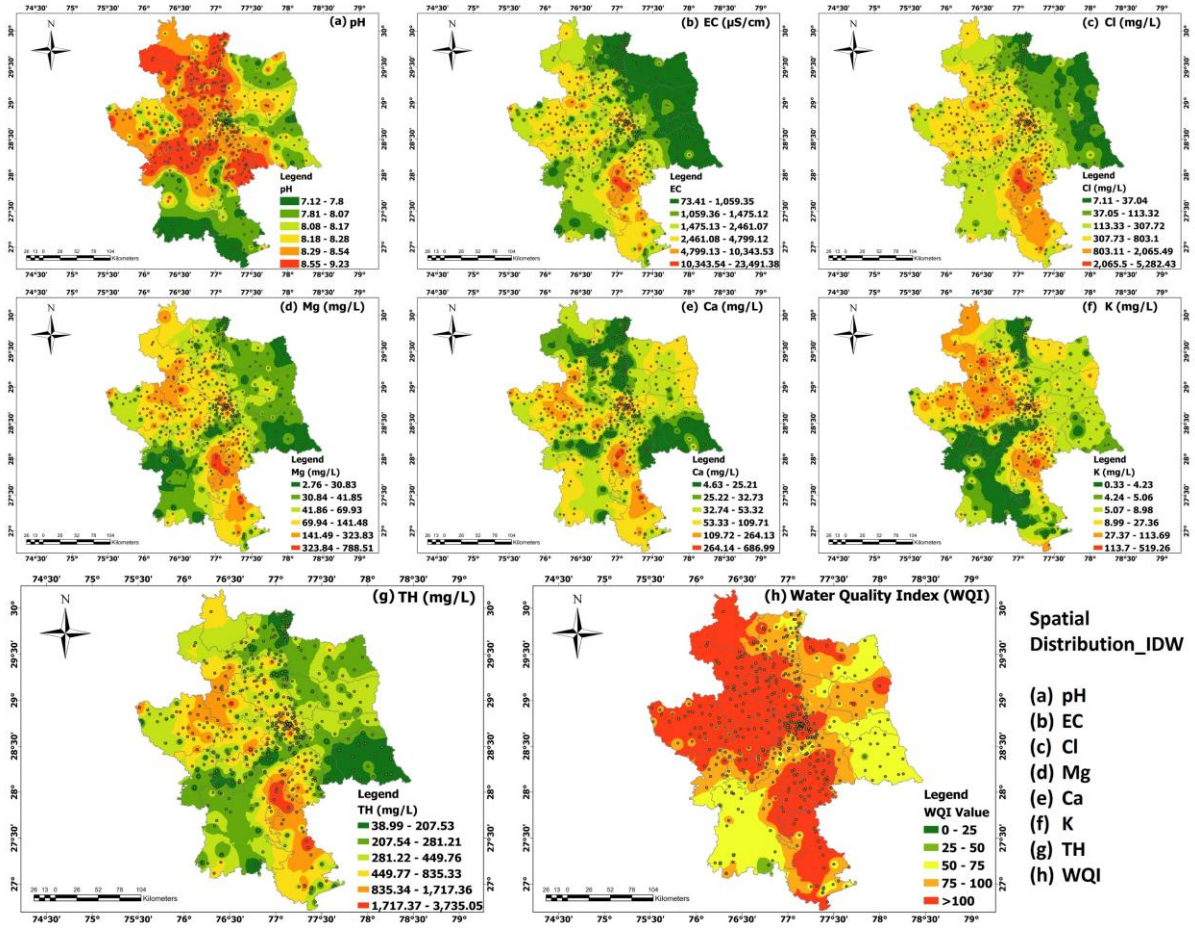


Figure 5.7 Spatial Distribution map showing geospatial variation of geochemical attributes (pH, EC, Cl, Mg, Ca, K, TH) and WQI in Delhi NCR, India.

The scaling ranges from excellent to unsuitable water category in different locations. The results revealed that in majority of the regions the WQI is > 100 suggest unsuitable water quality in the selected study region. Similar to red category the yellow category water quality is widely found in the study region followed by orange category. The entire study region is dominated by water quality found in poor, very poor and unsuitable for use category having indexing range 50 to 75, 75 to 100 and > 100 (Table 5.1). The extensive presence of high WQI values (dark red) suggests severe groundwater contamination, potentially due to industrial discharge, agricultural runoff, and leachate from waste disposal sites. The spatial distribution highlights localized pockets of better water quality, possibly influenced by natural recharge areas or lesser anthropogenic impact. The results identify the high-risk zones with severe contamination and zones with good water quality through spatial distribution of WQI parameters. The IDW maps clearly show that the groundwater's physicochemical properties and overall quality vary significantly according to their location. The central and southern zones have major contaminants and hardness however the border areas possess considerably safer and better water quality. Such spatial maps significantly help in area-specific water management and treatment planning. The Pearson correlation analysis and Geospatial distribution maps for all geochemical parameters of water quality show similar analysis across the study area.

5.3.3 PERFORMANCE ANALYSIS OF MACHINE LEARNING MODELS TO PREDICT WQI

The study employed six types of ML models, including Deep Learning model (ANN), Ensemble models (RF, XGBoost) and tree based (DT), Kernel-Based (SVM) and Instance-Based (KNN) to predict the Water Quality Index. The parameters used to predict WA-WQI are 'pH,' 'EC,' 'Cl,' 'Mg,' 'K,' 'TH,' and 'Ca,'. Table 2.4 compares ML approaches to forecast water quality from previous studies to the present study. Data preprocessing has been done initially on the study region dataset

to make it suitable for machine learning algorithms. Hyperparameter tuning has been conducted on each model following the completion of their initial training. The Nested K-Fold Cross Validation technique is used with k = 5. Nested cross-validation uses two stages, the outer cross-validation splits the data for testing, while the inner cross-validation tunes hyperparameters (Yang et al., 2021). The final model is trained on the full training set and tested on the outer set. K-fold cross-validation gives a more reliable performance estimate by averaging results from multiple splits, unlike a single train/test split (Singh et al., 2021). Grid Search hyperparameter optimization is applied to identify the optimal parameter combination with sufficient resources. Table 5.3 presented the model architecture obtained via hyperparameter tuning.

This study used datasets for training purposes through three-fold (K=3, K=4, K=5) cross-validation, dividing the dataset into 67%, 75%, and 80% for training, and 33%, 25%, and 20% for testing the models. This approach enhances accuracy and mitigates both underfitting and overfitting for unseen data, thereby improving the reliability of performance metrics. The six machine learning models used were effectively trained on the dataset across all folds, as illustrated in Figure 5.8. The ANN exhibits a significant reduction in MAPE error at 75% (K = 4), in comparison to all other models evaluated, indicating its sensitivity to data size.

The optimal performance of each machine learning model is achieved only when its hyperparameters are appropriately configured. This study has optimized the specific hyperparameters (Table 5.3) for each model, including 'max_depth' for Decision Trees, 'n_estimators' for Random Forests, and 'epochs' and 'optimizer' for ANN, to enhance data learning efficiency and mitigate overfitting. SVM model is configured with the following parameters: 'C': 0.1, 'gamma': 'scale', and 'kernel': 'linear'. These parameters define the margin width of the SVM and the selection of the kernel, which enhances the accuracy of WQI predictions. Hyperparameters used for the Decision Tree can grow to significant depths; however, the minimum necessary samples are established for leaf nodes and splits to maintain model balance. The selected

parameters for ANN hyperparameters help the model to learn and effectively capture nonlinear relationships in the dataset rapidly. Hyperparameters such as 'n_neighbors', 'learning_rate', and 'batch_size' directly influence the model's learning approach, predictive accuracy, and outcomes. These parameters facilitate a gradual training process with data diversity, resulting in robust WQI predictions.

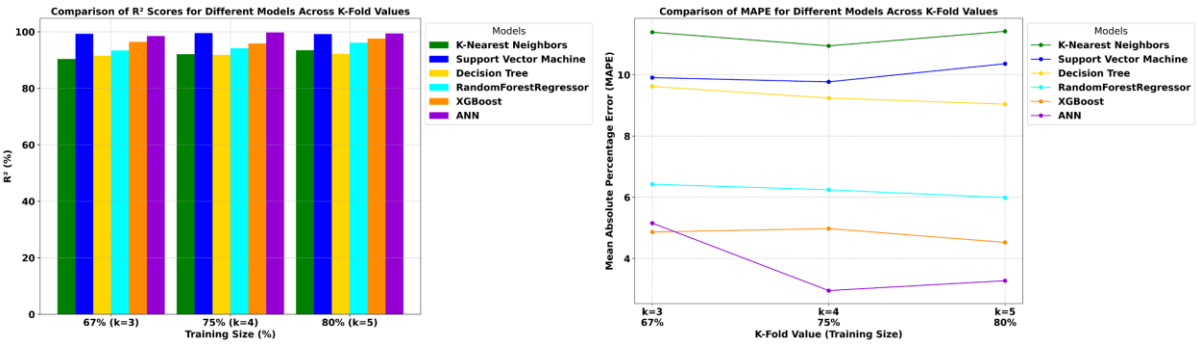


Figure 5.8 Comparison of R^2 scores and MAPE for various models in training over k-fold values

Table 5.3 Model architecture obtained via hyperparameter tuning (All 7 Features)

Model	Hyperparameters
K-Nearest Neighbors	{'n_neighbors': 3, 'p': 2, 'weights': 'distance'}
Support Vector Machine	{'C': 0.1, 'gamma': 'scale', 'kernel': 'linear'}
Decision Tree	{'max_depth': None, 'min_samples_leaf': 2, 'min_samples_split': 5}
Random Forest Regressor	{'max_depth': 10, 'min_samples_leaf': 2, 'min_samples_split': 5, 'n_estimators': 50}
XGBoost	{'learning_rate': 0.1, 'max_depth': 3, 'n_estimators': 600, 'subsample': 0.8}
ANN	{'batch_size': 10, 'epochs': 50, 'model__activation': 'relu', 'model__neurons': 48, 'model__optimizer': 'adam'}

5.3.4 ANALYSIS OF MODEL PERFORMANCE UTILIZING ALL 7 FEATURES BASED ON PEARSON CORRELATION COEFFICIENT TO PREDICT WQI

This study employs key water quality parameters 'pH,' 'EC,' 'Cl,' 'Mg,' 'K,' 'TH,' and 'Ca,' as features to predict the Water Quality Index (WQI) using the Nested K-Fold Cross Validation technique (k = 5) to enhance model reliability. Six machine learning models KNN, SVM, DT, RFR, XGBoost, and ANN as shown in Fig. 6. were evaluated using R^2, MAE, MAPE, RMSE, and Centered RMSE. ANN and SVM outperformed other models, achieving the highest R^2 (0.99), lowest MAE (4.518 for ANN), and minimal RMSE (8.294 for ANN, 9.689 for SVM), indicating superior predictive accuracy. XGBoost and Random Forest also demonstrated strong performance, with relatively low errors. Additionally, model ranking based on distance from the reference point in the Taylor diagram confirms ANN as the most accurate model (distance = 14.683), followed by SVM (15.146) and XGBoost (32.001). Random Forest, KNN, and Decision Tree showed comparatively lower accuracy, with DT ranking the lowest (57.001). The results highlight ANN and SVM as the most effective models for WQI prediction, offering high accuracy and minimal prediction errors. Based on this ANN has chosen for further validation.

The performance metrics (Table 5.4) used to evaluate the accuracy of predicted outcomes relative to actual values for all applied ML models are the coefficient of determination (R^2), mean absolute error (MAE), mean absolute percentage error (MAPE), root mean square error (RMSE) and centered root mean squared error (CRMSE). Figure 5.9 presents a comparison of each evaluation metrics, illustrating the model's performance in relation to various errors. R^2 reflects data variance, with values closer to 1 suggesting greater model performance. Lower MAPE percentages indicate better model accuracy. The ANN model demonstrates a prediction accuracy of 96.72% based on MAPE, and it exhibits an R^2 value of 0.99, indicating superior accuracy compared to all other applied machine learning models. The SVM demonstrates the second highest performance model

as R^2 (0.99) with RMSE (9.68). The XGBoost shows strong performance with R^2 (0.98) and RMSE (21.79), although it exhibits greater error compared to ANN and SVM. RF yields favorable results with R^2 (0.96) and RMSE (28.82) as compared to KNN, which has moderate performance with a maximum MAPE (11.42) and RMSE (37.52). The Decision Tree exhibits the lowest reliability compared to all other models, with the greatest RMSE of 42.972. ANN outperformed the other five models in this metrics comparison. The actual and predicted WQI for each model are presented in a line plot and a scatter plot as illustrated in Figure 5.10 The best-fit line and equation for each model are provided in the legend of each graph. ANN also performed well on the line plot and scatter plot of actual and predicted values, as most of the samples fell on the best-fit line, proving its accuracy.

Table 5.4 Performance metrics of ML models using all 7 features

Model	R^2	MAE	MAPE	RMSE	Centered RMSE
K-Nearest Neighbors	0.93	18.626	11.421	37.523	36.589
Support Vector Machine	0.99	9.029	10.362	9.689	4.948
Decision Tree	0.92	17.209	9.040	42.972	42.883
Random Forest Regressor	0.95	10.822	5.990	28.820	28.657
XGBoost	0.96	8.576	4.529	21.938	21.791
ANN	0.99	4.518	3.276	8.294	8.236

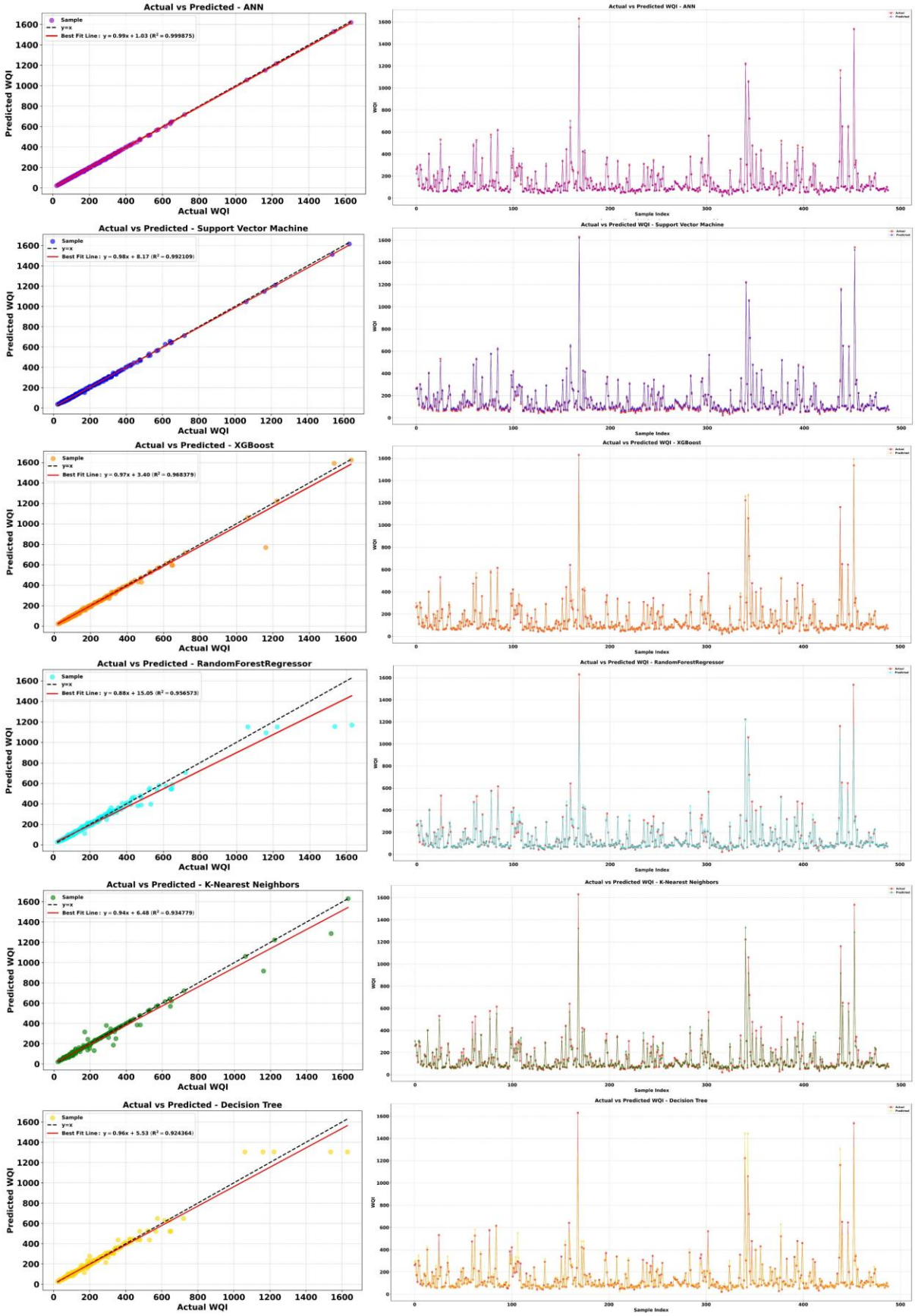


Figure 5.9 Scatter plot Actual Vs Predicted WQI for various ML models (a) ANN (b) SVM, (C) XGBoost (d) RF, (e) KNN (f) DT

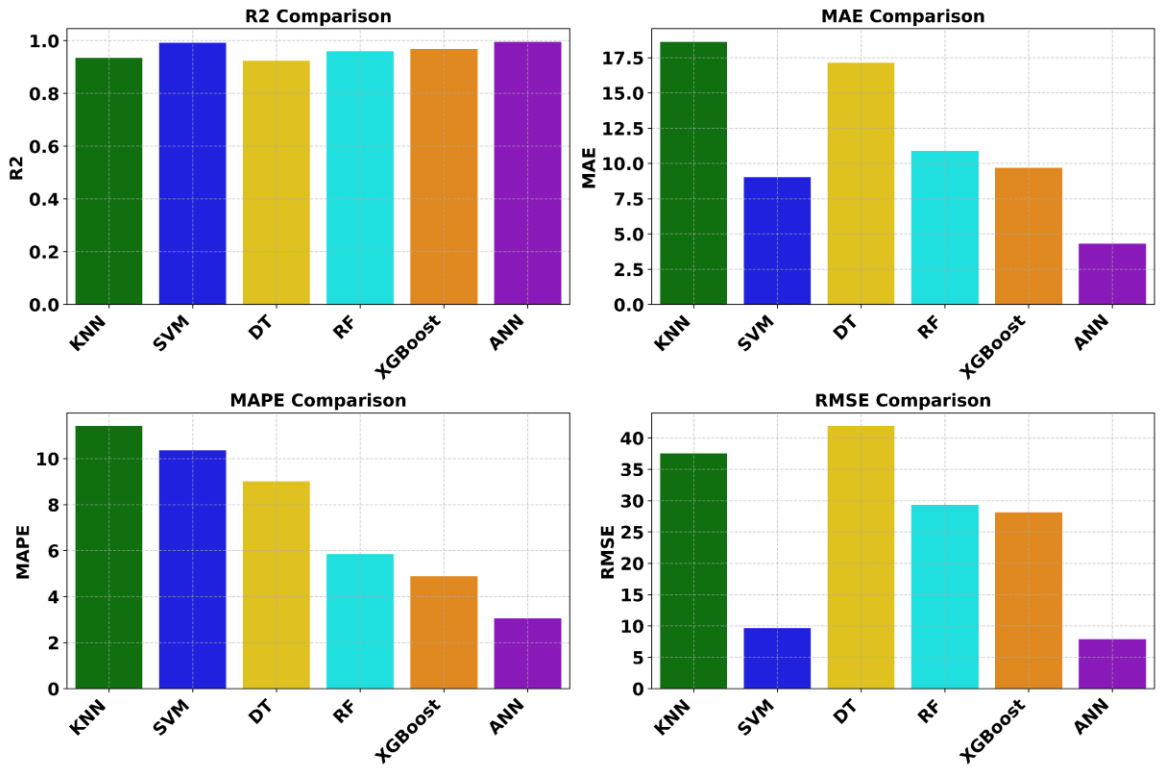


Figure 5.10 Bar graphs showing comparison among ML models on basis of different metrics

Taylor Diagram Analysis using all 7 features

Taylor Diagrams are polar plots that use standard deviation, Pearson's correlation coefficient, and centered root mean squared error to compare models to observed data. Model performance is assessed using the Taylor diagram (Figure 5.11), by measuring the distance of each model's point from a reference point (Javadi et al., 2022). The Centered RMSE is calculated on the basis of Equation 5.6 (Taylor, 2001). A model closest to the reference point has a standard deviation similar to the observed data, a high correlation, and a CRMSE close to zero, indicating better performance (Taylor, 2001). Based on the Taylor approach (Figure 5.11), the ANN has the lowest distance from the reference point, making it the best predicted model for the study region.

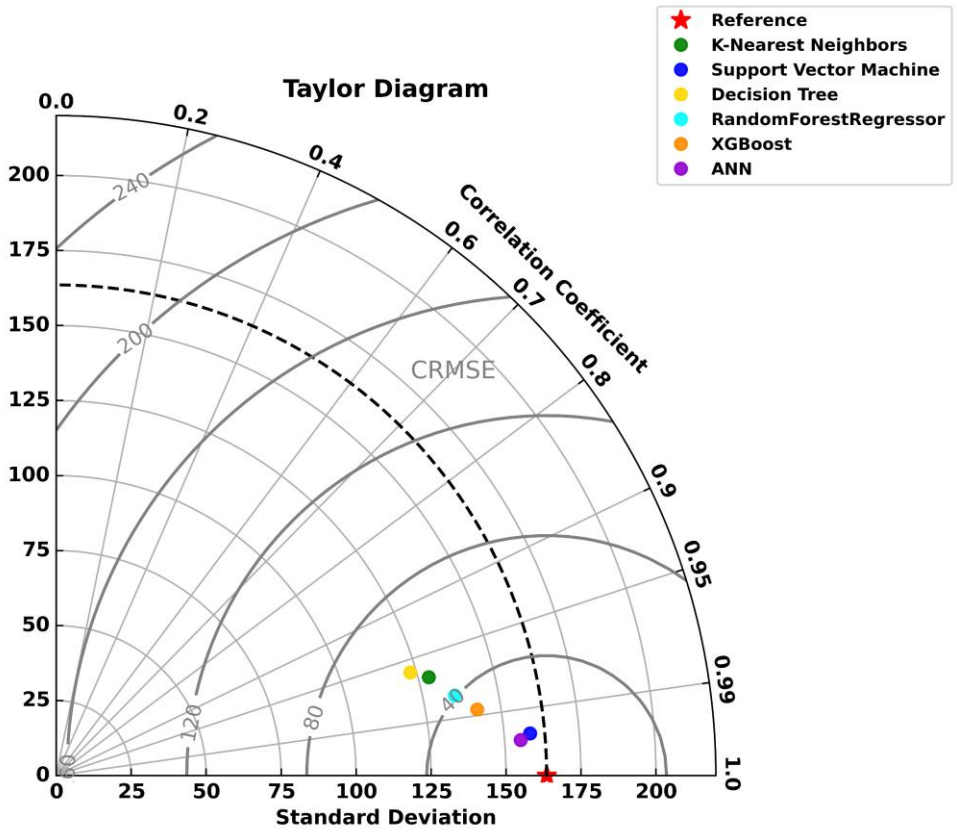


Figure 5.11 Taylor diagram using all 7 features

The Taylor technique ranks models by distance from the reference point: ANN ranks 1, SVM ranks 2, XGBoost ranks 3, Random Forest Regressor ranks 4, KNN ranks 5, and Decision Tree ranks 6 (Table 5.5). If all seven parameters used, ANN is the best predicted model, while DT is less accurate for the study region. Consequently, ANN was selected for further validation using a separate dataset of 70 samples.

Table 5.5 Models Ranked on basis of distance from reference point in Taylor diagram

Rank	Model	Distance from Reference
1	ANN	14.683
2	Support Vector Machine	15.146
3	XGBoost	32.001
4	RandomForestRegressor	40.635
5	K-Nearest Neighbors	51.201
6	Decision Tree	57.001

5.3.5 ANALYSIS OF MODEL PERFORMANCE USING TOP 4 FEATURES BASED ON PEARSON CORRELATION COEFFICIENT TO PREDICT WQI

The features used to predict WQI are EC, Mg, K, TH, and the target variable is WQI. The Nested K-Fold Cross Validation technique is used with k = 5. Table 5.6 shows the model architecture obtained via hyperparameter tuning for top 4 features. Figure 5.12 (a-f) shows the line plot of actual vs. predicted WQI for KNN, SVM, DT, RF, XGBoost, ANN models for top 4 features.

Table 5.6 Performance Metrics of ML Models Using Top 4 Features (EC, Mg, K, TH)

Model	R^2	MAE	MAPE	RMSE	Centered RMSE
K-Nearest Neighbors	0.96	16.440	11.298	31.773	31.162
Support Vector Machine	0.99	9.280	9.817	11.112	10.759
Decision Tree	0.92	18.634	12.781	42.603	42.413
Random Forest Regressor	0.96	12.876	9.452	29.183	28.954
XGBoost	0.98	12.022	9.657	21.731	21.621
ANN	0.99	8.062	8.444	10.499	10.432

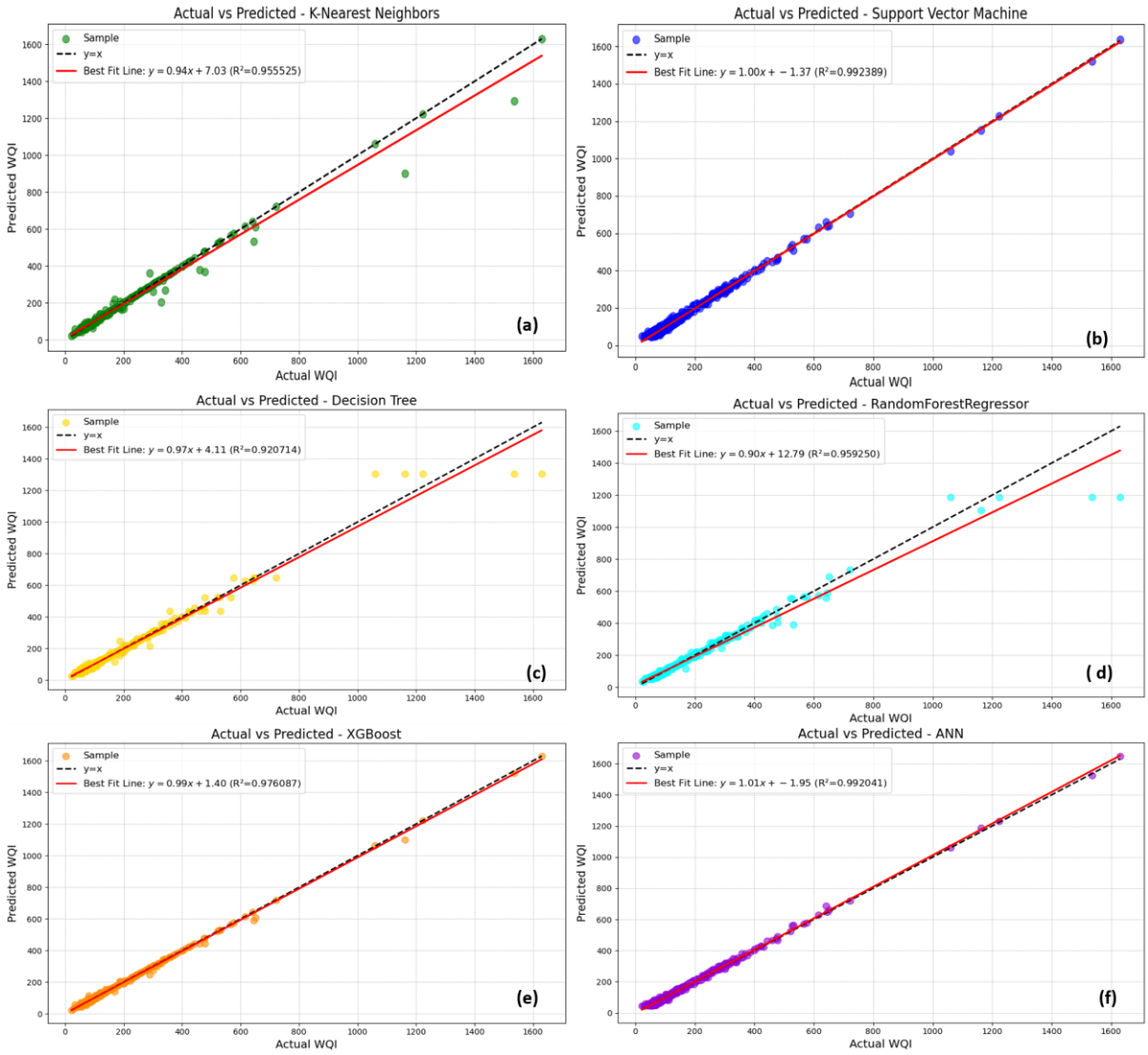


Figure 5.12 Line plot of Actual vs. Predicted WQI for (a) KNN, (b) SVM, (c)DT, (d)RF, (e)XGBOOST, (f)ANN using top 4 features

Taylor Diagram Analysis using top 4 features

The Taylor diagram in Figure 5.13 provides a visual representation of the performance of machine learning models using the top four features. Based on the rankings in Table 5.7, which evaluates models by their distances from the reference point in the Taylor diagram, the Support Vector Machine (SVM) model is identified as the most suitable. Consequently, this model is selected for further validation.

Table 5.7 Models Ranked on Basis of Distance from Reference Point in Taylor Diagram (Top 4 Features)

Rank	Model	Distance from Reference
1	Support Vector Machine	17.477
2	ANN	17.541
3	XGBoost	31.848
4	Random Forest Regressor	41.332
5	K-Nearest Neighbors	43.570
6	Decision Tree	56.949

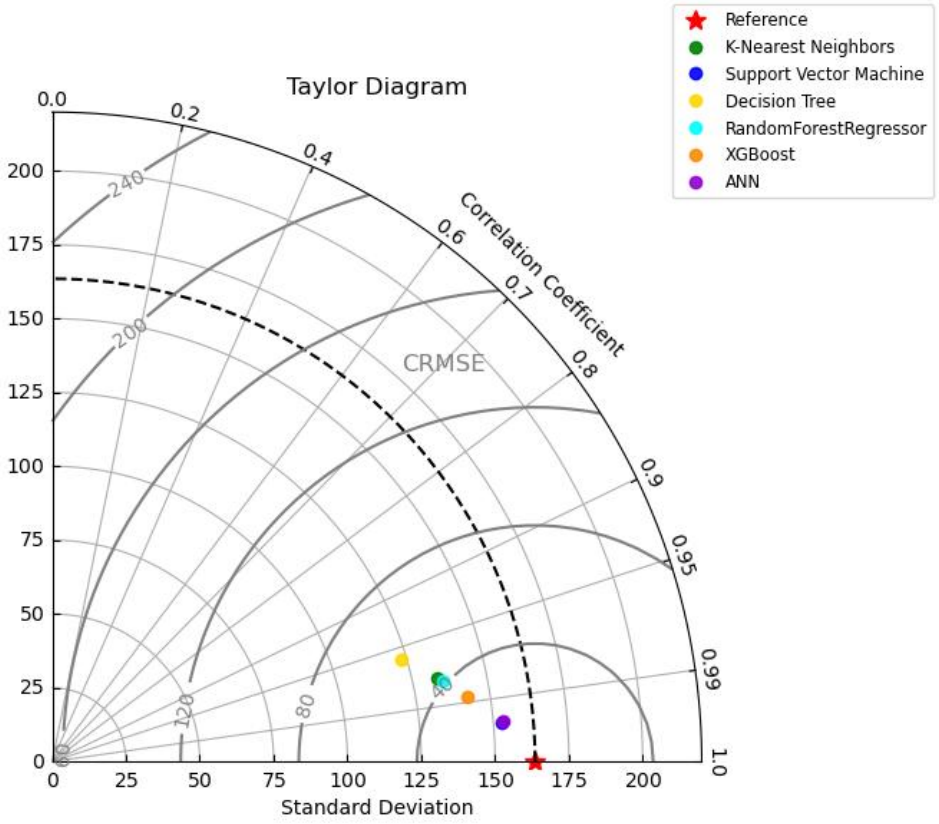


Figure 5.13 Taylor diagram (using top 4 features)

5.3.6 ANALYSIS OF MODEL PERFORMANCE USING TOP 2 FEATURES BASED ON PEARSON CORRELATION COEFFICIENT TO PREDICT WQI

WQI as the target variable. The Nested K-Fold Cross-Validation technique is applied with k = 5 to ensure robust model evaluation. The performance metrics of various machine learning models for these features are detailed in Table 5.8, while Figure 5.14 depicts a line plot showing actual vs. predicted WQI values for KNN, SVM, DT, RF, XGBoost, and ANN models.

Table 5.8 Performance Metrics of ML Models Using Top 2 Features (Mg, K)

Model	R^2	MAE	MAPE	RMSE	Centered RMSE
K-Nearest Neighbors	0.96	15.510	11.980	30.726	30.359
Support Vector Machine	0.99	11.821	10.733	15.527	15.060
Decision Tree	0.93	18.914	13.204	38.194	38.007
Random Forest Regressor	0.95	14.975	11.017	32.731	32.478
XGBoost	0.97	14.420	11.954	25.671	25.489
ANN	0.99	11.236	10.161	15.815	15.491

Taylor Diagram Analysis using top 2 features

The Taylor diagram in Figure 5.15 illustrates the performance of ML models using the top two features. Table 5.9 ranks the models based on their distances from the reference point in the Taylor diagram, considering the top two features. Based on these rankings, the SVM model with the top two features is selected for further validation.

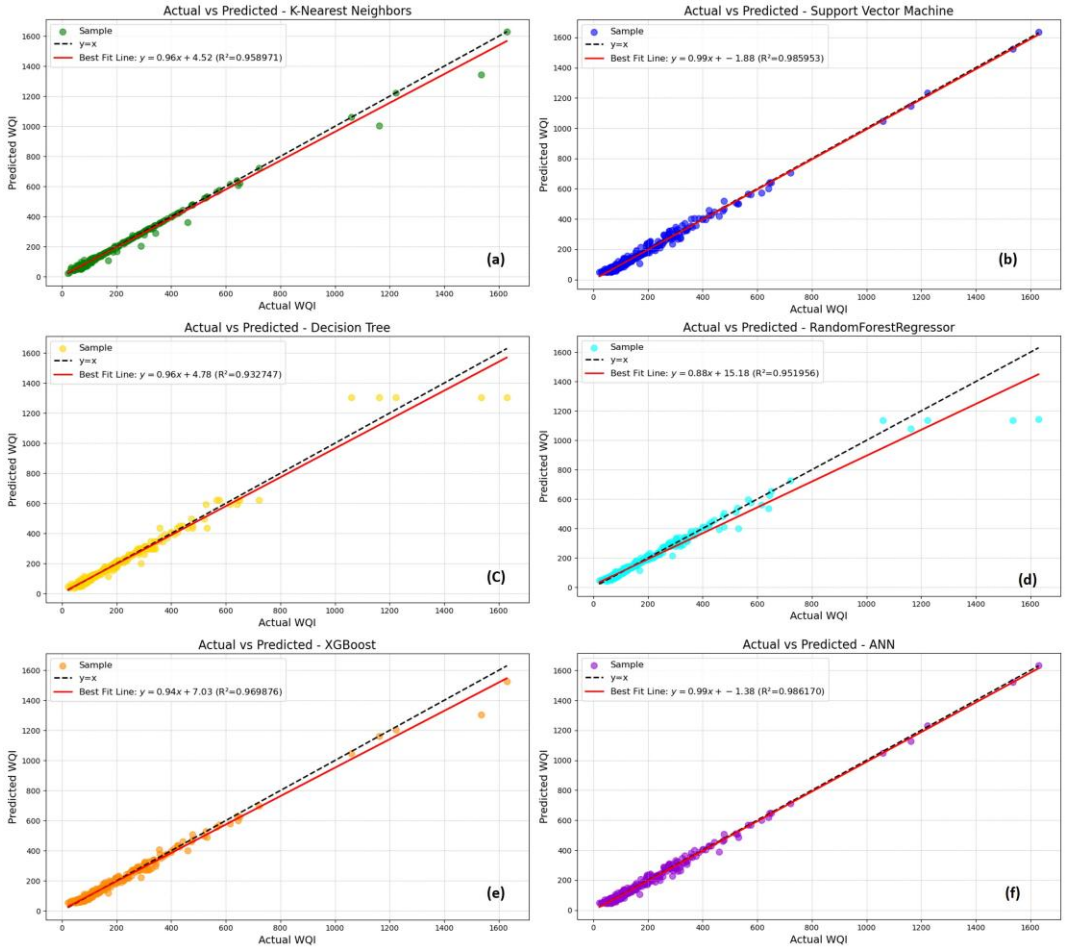


Figure 5.14 Line plot of Actual vs. Predicted WQI for (a) KNN, (b)SVM, (c)DT, (d)RF, (e)XGBOOST, (f)ANN using top 2 features

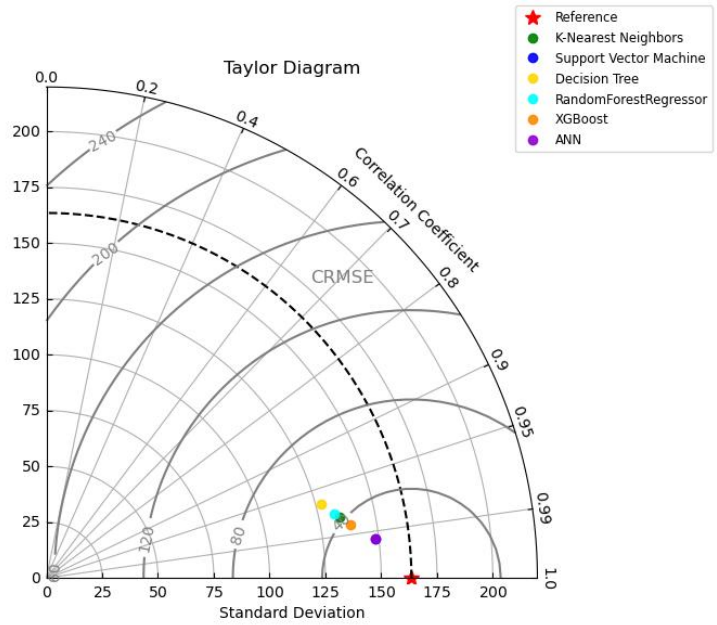


Figure 5.15 Taylor diagram (using top 2 features.)

Table 5.9 Models Ranked on Basis of Distance from Reference Point in Taylor Diagram (Top 2 Features)

Rank	Model	Distance from Reference
1	Support Vector Machine	23.789
2	ANN	23.931
3	XGBoost	36.253
4	K-Nearest Neighbors	42.236
5	Random Forest Regressor	45.194
6	Decision Tree	52.376

5.3.7 ANALYSIS OF MODEL PERFORMANCE USING TOP MOST FEATURE BASED ON PEARSON CORRELATION COEFFICIENT TO PREDICT WQI

The top most feature used for prediction is 'K' and the target variable is WQI. The Nested K-Fold Cross Validation technique is used with k = 5. Table 5.10 displays the performance metrics of various machine learning models for this feature. Consequently, Fig. 5.16 (a-f). depicts a line plot of actual vs. predicted WQI for KNN, SVM, DT, RF, XGBoost, and ANN models using only the most significant feature.

Table 5.10 Performance Metrics of ML Models Using Single Feature (K)

Model	R^2	MAE	MAPE	RMSE	Centered RMSE
K-Nearest Neighbors	0.761	39.069	28.492	64.739	64.464
Support Vector Machine	0.812	34.187	24.920	57.326	55.596
Decision Tree	0.746	39.331	27.851	69.024	68.802
Random Forest Regressor	0.784	37.460	27.064	63.456	63.174
XGBoost	0.813	35.829	26.425	59.269	58.981
ANN	0.840	34.297	27.011	52.606	52.239

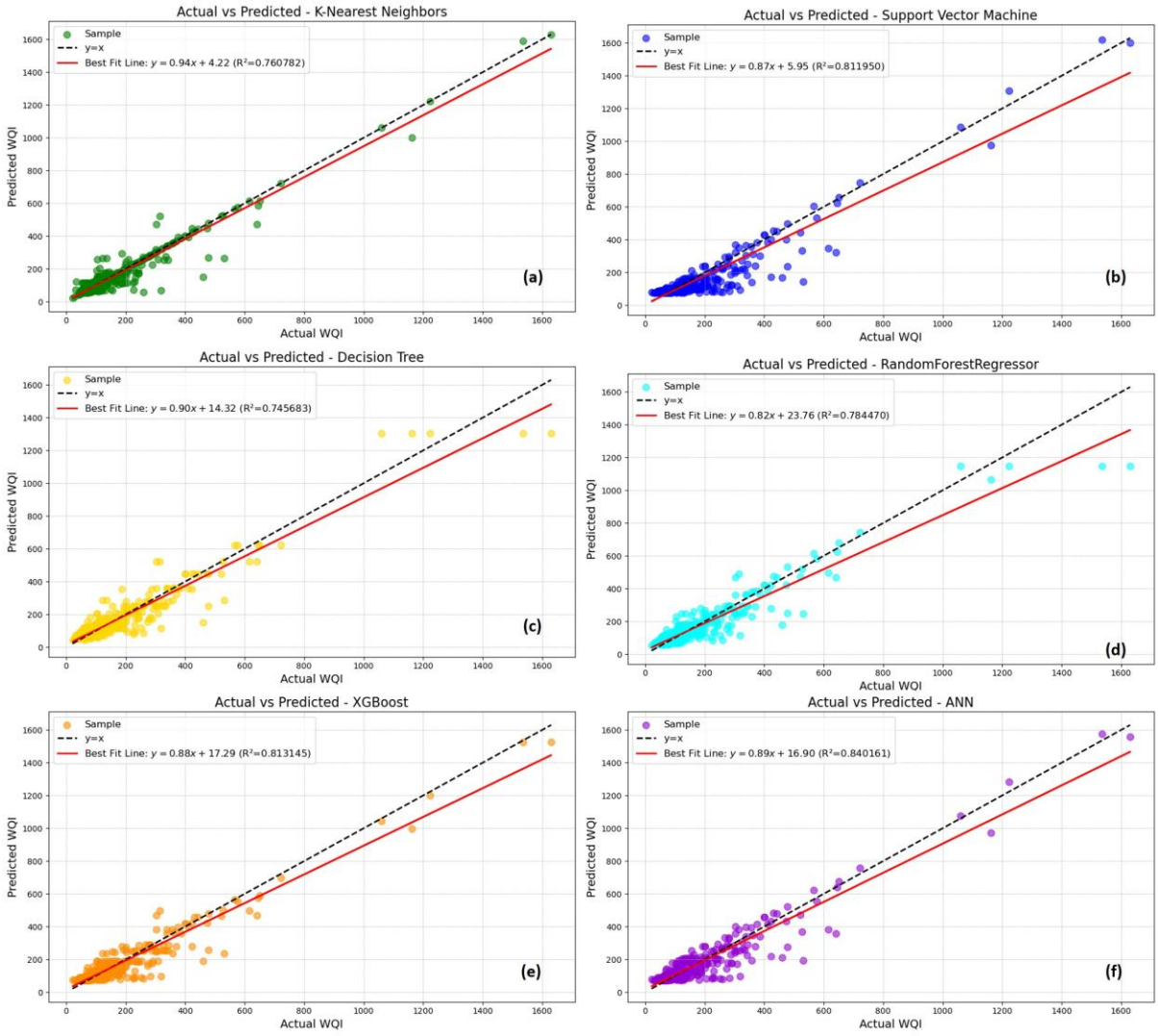


Figure 5.16 Line plot of Actual vs. Predicted WQI for (a)KNN, (b)SVM, (C)DT, (d)RF, (e)XGBOOST, (f)ANN using top most features (K)

Taylor Diagram Analysis using top most feature

The Taylor diagram in Figure 5.17 illustrates the performance of machine learning models using the most significant feature. Table 5.11 ranks the models based on their distances from the reference point in the Taylor diagram when using this single top feature. Based on the rankings, the ANN model, utilizing only the most significant feature, is selected for further validation.

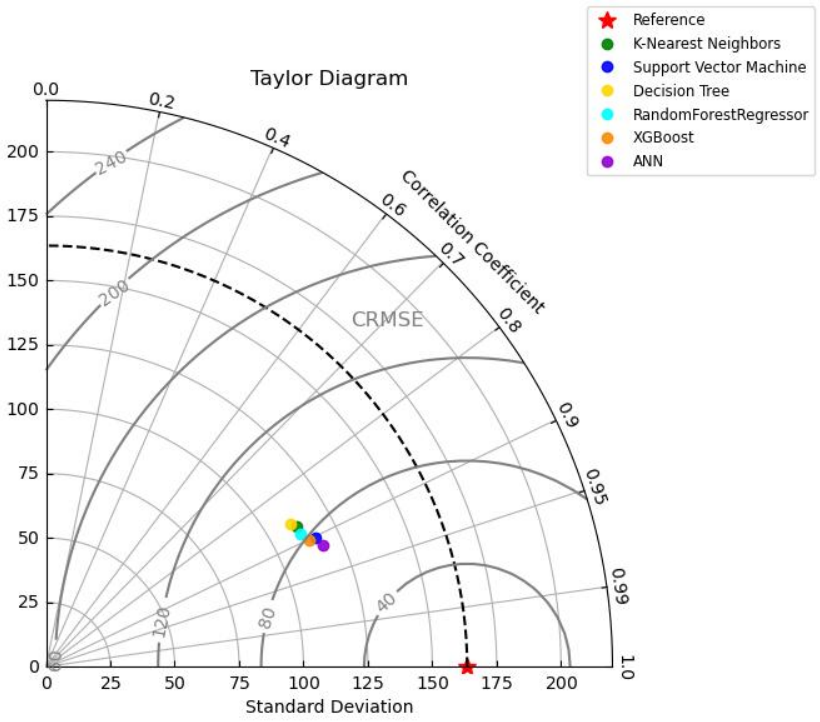


Figure 5.17 Taylor diagram (using top most feature 'K')

Table 5.11 Models Ranked on Basis of Distance from Reference Point in Taylor Diagram (Single Feature K)

Rank	Model	Distance from Reference
1	ANN	72.9637
2	Support Vector Machine	77.559
3	XGBoost	78.541
4	Random Forest Regressor	82.9255
5	K-Nearest Neighbors	85.7171
6	Decision Tree	88.2782

5.3.8 FURTHER PERFORMANCE CHECK AND VALIDATION ON THE BEST ML MODELS

For each of the four cases with different input feature sets, a model was selected using the Taylor diagram for further prediction testing on an independent dataset of 70 samples. Table 5.12 presents

the best-performing models for all four cases, while Table 5.13 provides the performance metrics of four. Figures 5.18, 5.19, 5.20, 5.21 display scatter plots comparing predicted and actual WQI values for ANN (Case-1), Support Vector Machine (Case-2), Support Vector Machine (Case-3), and ANN (Case-4). Additionally, Figure 5.22 illustrates the Taylor diagram for further validation and performance assessment of all applied models, with Table 5.14 ranking the models based on their distance from the reference point in the diagram. From Tables 5.12 and 5.14, it is evident that models with a higher number of input features generally exhibited better performance. ANN (Case-1) emerged as the top-performing model; however, SVM (Case-3), which utilized only two features (Mg, K), demonstrated a very slight decrease in performance compared to ANN (Case-1), which required all seven features. Therefore, to optimize water quality monitoring, SVM (Case-3) is a practical choice, as it requires only two features, reducing both cost and time associated with laboratory testing of additional parameters.

Table 5.12 Best Performing Models in Different Feature Cases

Case	Best Model	Input Features
1	ANN	pH, EC, Mg, K, TH, Ca, Cl
2	Support Vector Machine	EC, Mg, K, TH
3	Support Vector Machine	Mg, K
4	ANN	K

Table 5.13 Performance Metrics of Validated Models on Independent 70-Sample Dataset

Model (Case)	R^2	MAE	MAPE	RMSE	Centered RMSE
ANN — Case 1 (All 7 features)	0.999	3.746	2.550	5.520	5.373
SVM — Case 2 (Top 4: EC, Mg, K, TH)	0.998	6.510	5.526	8.227	7.626

Model (Case)	R^2	MAE	MAPE	RMSE	Centered RMSE
SVM — Case 3 (Top 2: Mg, K)	0.998	6.841	5.601	9.588	9.585
ANN — Case 4 (Single feature: K)	0.932	33.369	23.210	53.087	51.699

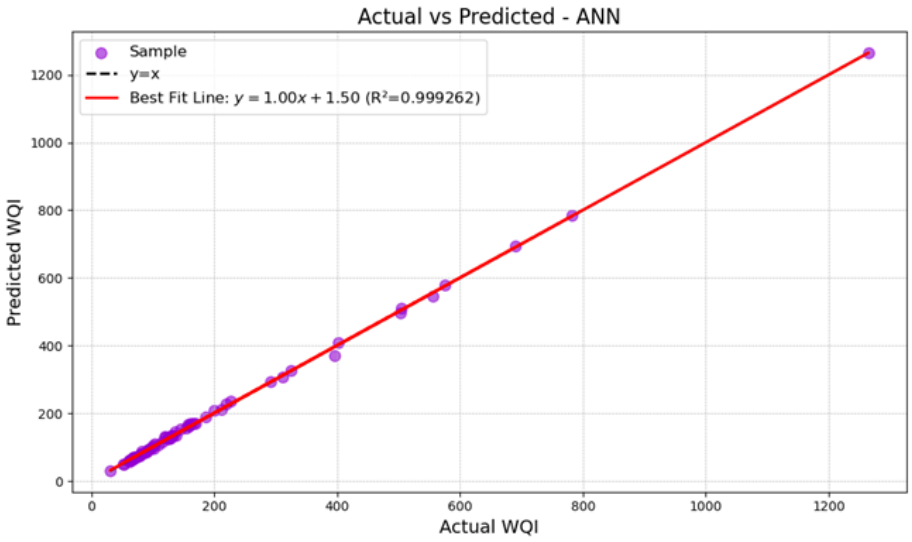


Figure 5.18 Scatter plot of Predicted vs Actual WQI of ANN (Case-1)

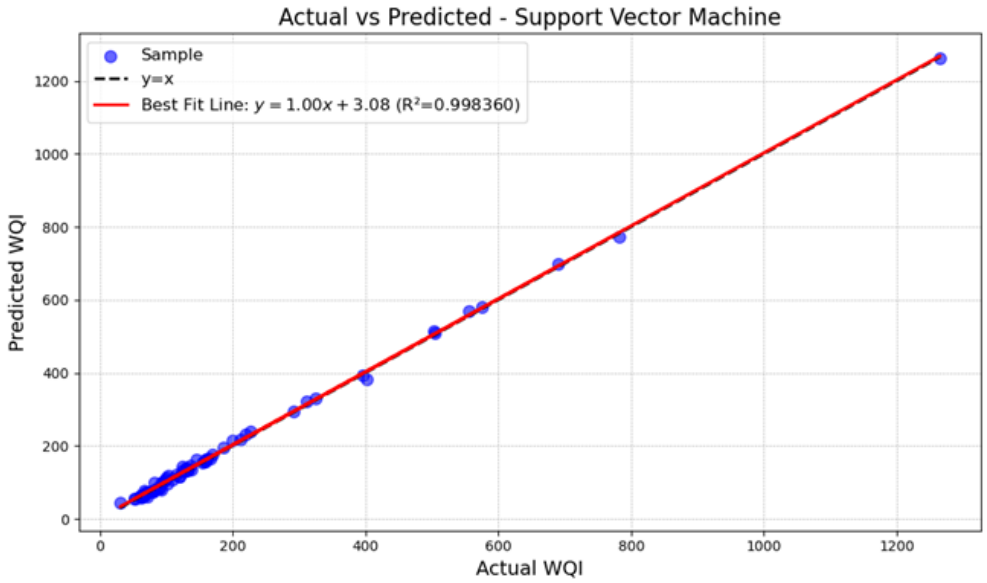


Figure 5.19 Scatter plot of Predicted vs Actual WQI of SVM (Case-2)

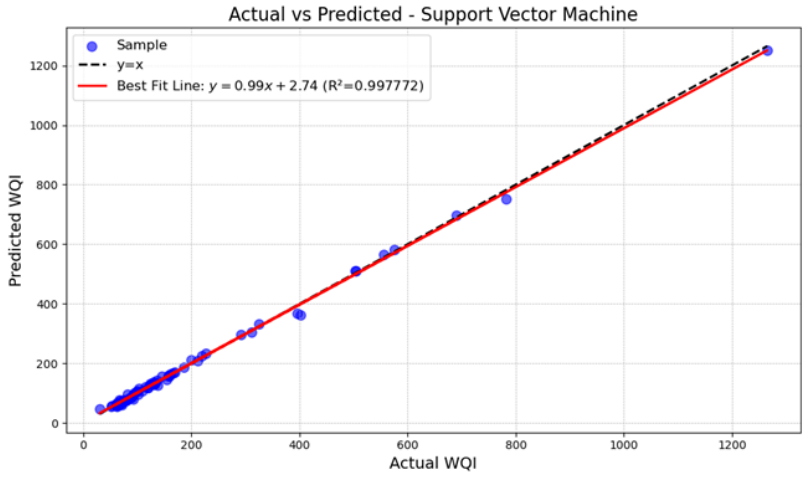


Figure 5.20 Scatter plot of Predicted vs Actual WQI of SVM (Case-3)

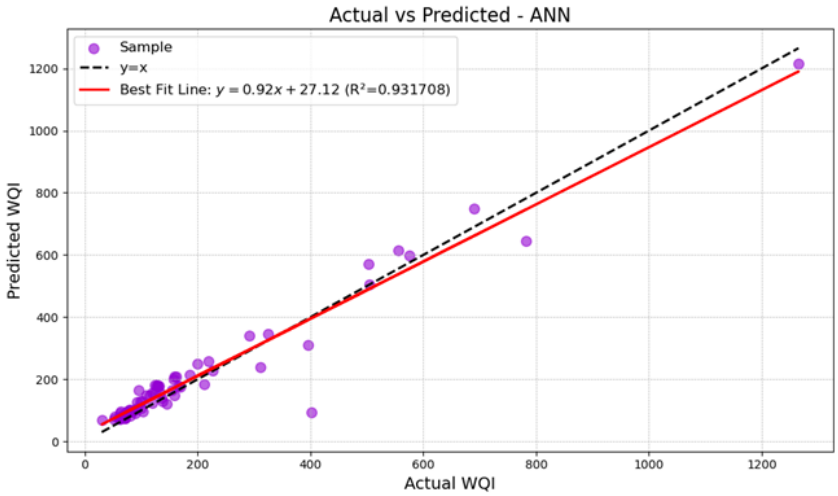


Figure 5.21 Scatter plot of Predicted vs Actual WQI of SVM (Case-4)

Table 5.14 Models Ranked on Basis of Distance from Reference Point in Taylor Diagram Final Validation

Rank	Model	Input Features	Distance from Reference
1	ANN (Case 1)	pH, EC, Mg, K, TH, Ca, Cl	7.6777
2	Support Vector Machine (Case 2)	EC, Mg, K, TH	11.1445
3	Support Vector Machine (Case 3)	Mg, K	13.4403
4	ANN (Case 4)	K	68.1224

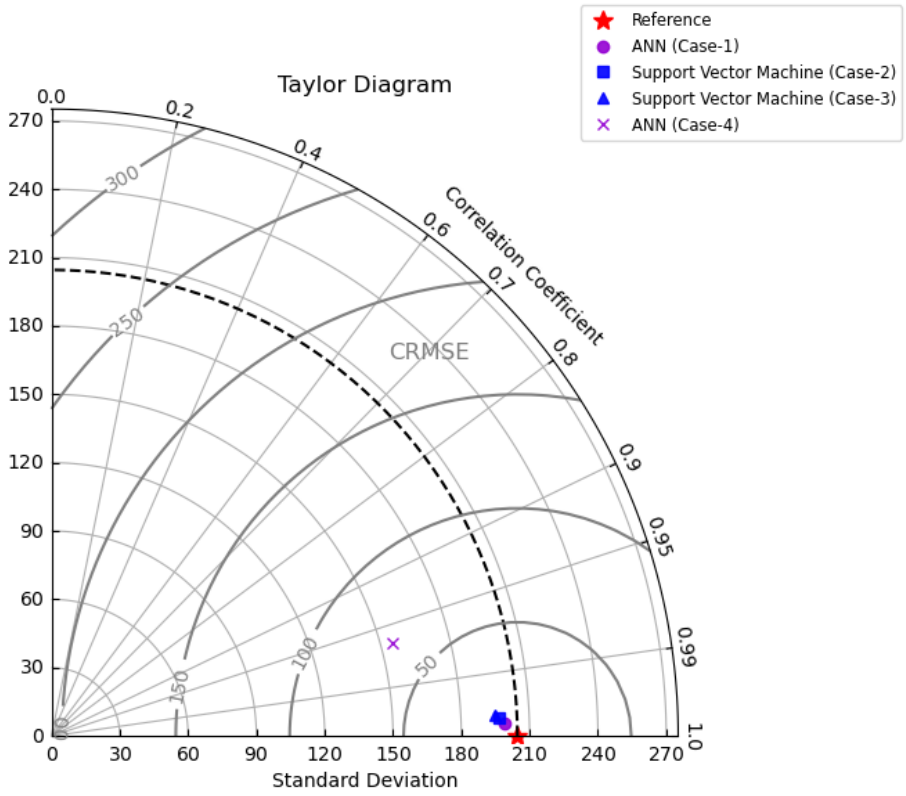


Figure 5.22 Taylor diagram for further performance check

5.3.9 ASSESSMENT OF SOIL PHYSICOCHEMICAL PROPERTIES AND HEAVY METAL CONTAMINATION

The soil contamination assessment was performed using 150 soil samples collected from 10 borehole locations around the Bhalswa landfill. The borehole depths ranged from 1 m to 15 m, while the distance from the reference point ranged from 306.81 m to 2012.93 m. The analysed soil parameters included pH, electrical conductivity, total dissolved solids, alkalinity, chlorides, total hardness, magnesium, aluminium, manganese, potassium, iron, zinc, copper, cadmium, chromium, and lead.

The soil analysis revealed significant vertical and horizontal variation in physicochemical characteristics and heavy metal concentrations. Heavy metal concentrations, including Fe, Zn, Cu,

Cd, Cr, and Pb, showed an increasing trend with depth. The observed trend indicates downward migration of pollutants through landfill leachate infiltration. Boreholes located closer to the landfill exhibited higher concentrations of heavy metals and dissolved solids, confirming the influence of distance from the contamination source. The results also showed that alkalinity and total hardness decreased with depth, while chlorides and total dissolved solids showed variable trends across boreholes.

Table 1 provides a comprehensive study of the physico-chemical statistics summary of all collected soil samples. For soil, the statistical distribution showed measurable contamination across all boreholes. The mean pH value was 7.5883, indicating slightly alkaline soil conditions. Electrical Conductivity ranged from 81.17 uS/cm to 713.70 uS/cm, while TDS ranged from 40.27 ppm to 350.20 ppm. The mean depth was 8 m, and the wide distance range confirmed spatial coverage around the landfill.

Table 5.15 Statistical summary of soil physicochemical properties and heavy metals

Parameter	Count	Mean	Standard Deviation	Minimum	Maximum
Distance	150	1168.033	560.5472	306.81	2012.93
Depth	150	8	4.335	1	15
pH	150	7.5883	0.4541	6.3	8.22
EC (uS/cm)	150	256.4929	130.8146	81.17	713.7
TDS (ppm)	150	126.6048	64.299	40.27	350.2
Alkalinity (mg/l)	150	84.8833	25.0933	50	190
Chlorides (mg/l)	150	42.0557	22.7714	14.98	114.82
Total Hardness (mg/l)	150	87.5407	31.1224	43.7	205
Mg (ppm)	150	14.8251	2.335	6.5	19.86
Al (ppm)	150	70.6644	46.2977	0	195.1
Mn (ppm)	150	1.9953	1.0359	0.4	5.1
K (ppm)	150	15.2024	11.4544	2.83	54.26
Fe (ppm)	150	125.7074	64.3051	30.92	338.54
Zn (ppm)	150	0.4113	0.1717	0.1	1.02
Cu (ppm)	150	0.0952	0.0696	0.01	0.43
Cd (ppm)	150	0.0246	0.0165	0.01	0.09
Cr (ppm)	150	0.1838	0.1117	0.02	0.56
Pb (ppm)	150	0.2257	0.1741	0.06	0.94

The heavy metal statistics showed marked variation across Fe, Al, Pb, Cr, and Cd, indicating that contamination intensity changed with depth and distance from the landfill. The soil data therefore supported the application of machine learning models for heavy metal contamination prediction.

The highest mean concentration among measured heavy metals was observed for Iron, followed by aluminium, manganese, zinc, lead, chromium, copper, and cadmium. The wide range of Fe, Al, Pb, Cr, and Cd values indicates unequal contamination intensity across boreholes and depths. The combined variation in depth, distance, physicochemical parameters, and heavy metals confirms that soil contamination around the landfill cannot be explained through a single parameter. Therefore, predictive modelling was required to capture nonlinear and multidimensional relationships among the measured variables.

5.3.10 PERFORMANCE ANALYSIS OF MACHINE LEARNING MODELS FOR SOIL SAMPLES

For soil heavy metal contamination prediction, RF-GA was compared with Random Forest, XGBoost, SVM, ANN, and Logistic Regression. The models were trained using multiple split ratios of 60%, 70%, 80%, and 90%. At 80% training data, RF-GA achieved accuracy of 94.00%, while RF achieved 92.17%, XGBoost achieved 91.67%, ANN achieved 90.00%, SVM achieved 89.76%, and Logistic Regression achieved 87.50%. The RF-GA model also achieved the highest F1-score of 0.94 at the same training level. The superior performance of RF-GA was attributed to Genetic Algorithm-based feature optimization, which reduced multicollinearity and filtered less relevant inputs. Figure 5.23 (a–g) illustrates that to guarantee a thorough assessment, each model was trained utilizing various data split ratios (60%, 70%, 80%, and 90%). Table 5.16 provides the analysis of the proposed RF+GA system in comparison to the recommended model, encompassing scenarios both with and without enriched feature section attributes.

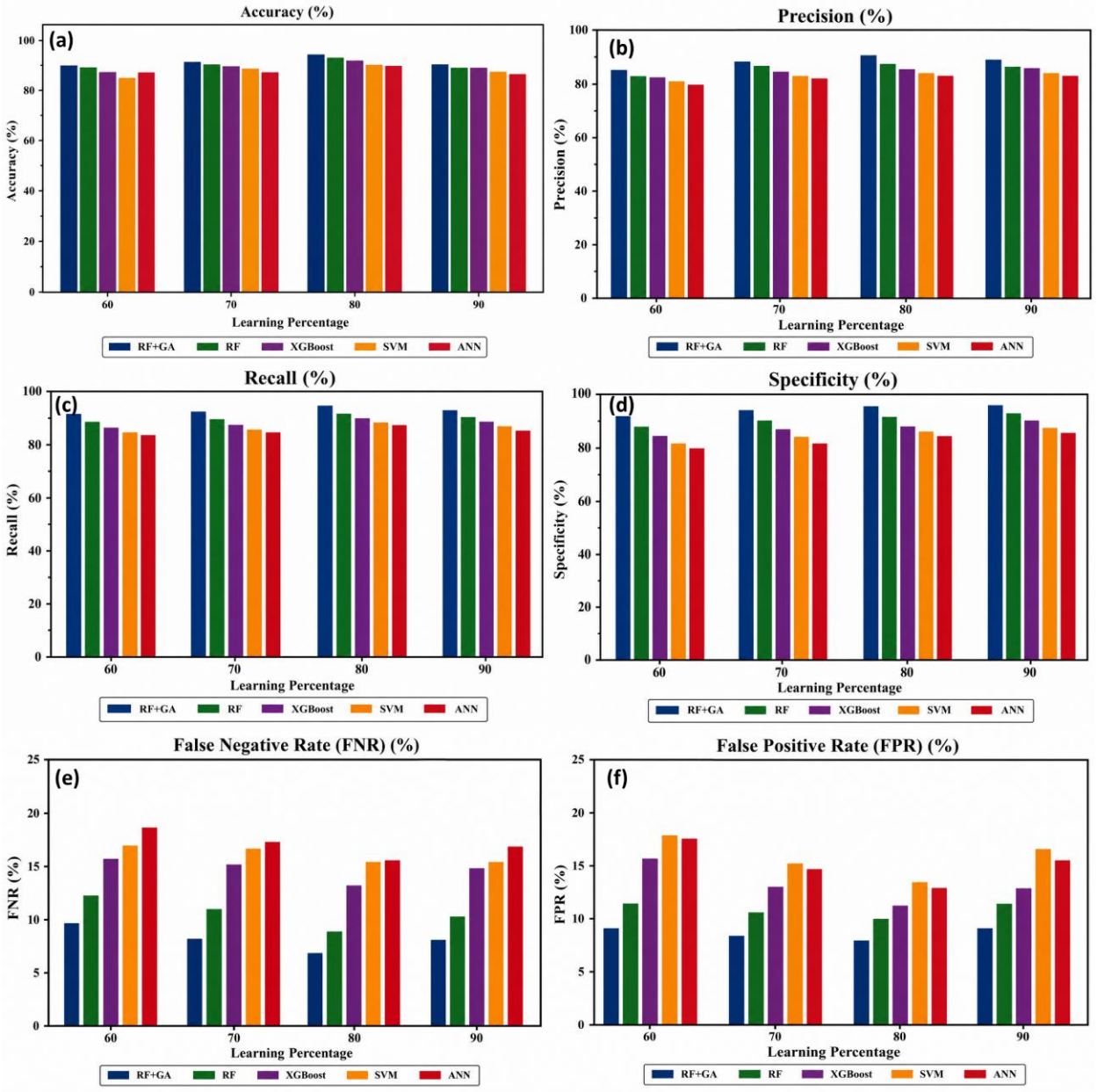


Figure 5.23 (a)-(g) Illustrates various model evaluation analysis using developed approach over basic models, (a) Accuracy (%), (b) Precision (%), (c) Recall (%), (d) Specificity (%), (e) False Negative Rate (%), (f) False Positive Rate (%)

5.3.11 MODEL VALIDATION FOR SOIL SAMPLES

For soil, model validation was reflected through comparative performance across training ratios and feature selection conditions. RF-GA with correlation-informed Chi-square feature selection achieved the best results across accuracy, recall, precision, sensitivity, specificity, False Negative Rate, and False Positive Rate. Accuracy increased from 92.978% for RF to 94.003% for RF-GA

with correlation features. Recall increased from 89.888% to 91.798%, precision increased from 88.170% to 91.035%, sensitivity increased from 90.227% to 93.095%, and specificity increased from 89.059% to 91.924%. The RF-GA model with correlation features showed the lowest False Negative Rate and False Positive Rate. False

Negative Rate decreased from 10.888% for RF to 6.601% for RF-GA with correlation features. False Positive Rate decreased from 9.998% to 7.974%. These values indicate improved classification reliability and lower misclassification risk for heavy metal contamination prediction.

Table 5.16 Soil model comparison with feature selection and optimization

Metrics	RF	RF + GA without correlation features	RF + GA with correlation features (Chi-square feature section)
Accuracy (%)	92.978	93. 3051	94.003
Recall (%)	89.888	90.457	91.798
Precision (%)	88.17	89.499	91.035
Sensitivity (%)	90.227	91.765	93.095
Specificity (%)	89.059	90.111	91.924
FNR (%)	10.888	8.288	6.601
FPR (%)	9.998	8.333	7.974

5.3.12 CONVERGENCE ANALYSIS OF THE OPTIMIZED MODEL

Convergence analysis was used to evaluate the optimization behaviour of the RF-GA model for soil contamination prediction. The convergence graph (Figure 5.24) compared RF-GA with RF, XGBoost, SVM, ANN, and Logistic Regression across iterations from 0 to 50. The y-axis represented cost function value, which reflected model error. Lower cost function values represented better model performance.

The RF-GA model demonstrated the most rapid and substantial reduction in cost function values. A significant reduction was observed around the 15th iteration, followed by consistently lower values between the 40th and 50th iterations. At the 40th iteration (Table 5.17), RF-GA reached a cost value of approximately 1.601%, outperforming RF at 1.921%, XGBoost at 2.032%, SVM at 2.184%, ANN at 2.413%, and Logistic Regression at 2.635%.

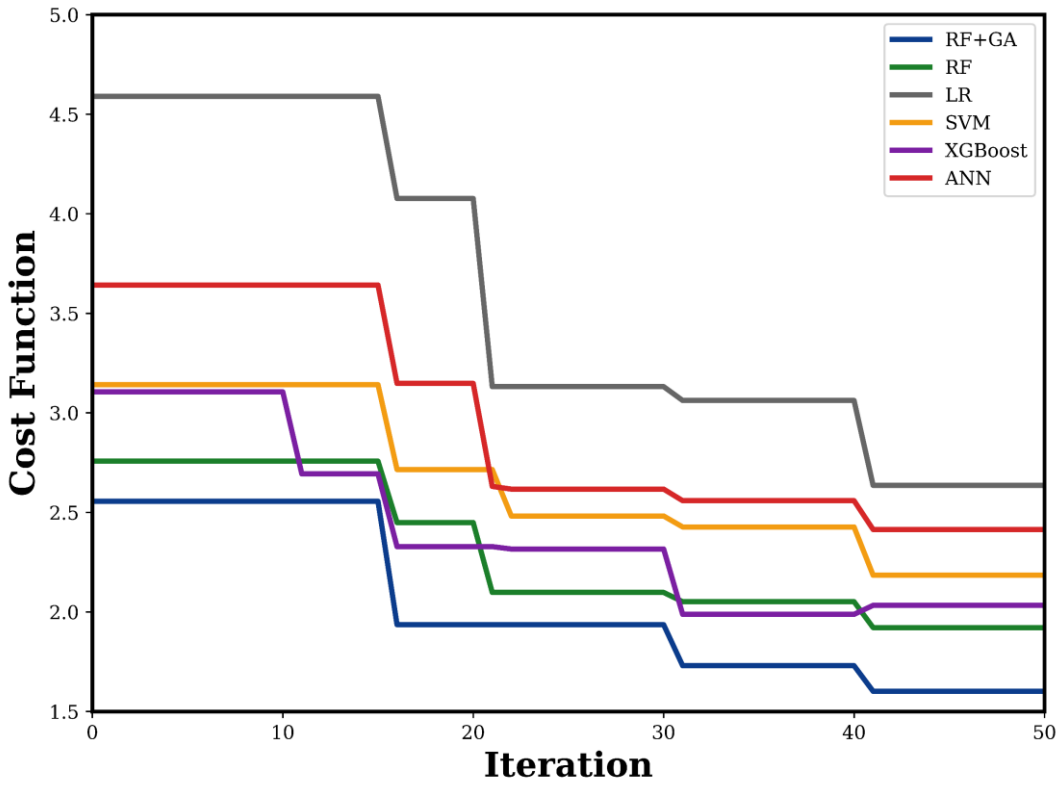


Figure 5.24 Illustrates convergence analysis of developed approach (RF-GA) over compared approaches (RF, LR, SVM, XGBoost, ANN)

The convergence results confirm the optimization advantage of RF-GA over the comparison models. RF and XGBoost also showed improvement, but both models failed to achieve the same level of cost reduction as RF-GA. SVM, ANN, and Logistic Regression showed weaker convergence behaviour, with Logistic Regression recording the highest cost function value among the compared models. The integration of Genetic Algorithm with Random Forest therefore improved convergence rate, reduced model error, and strengthened predictive performance for soil heavy metal contamination.

Table 5.17 Cost function values at the 40th iteration

Model	Cost function value (%)
RF + GA	1.601
Random Forest	1.921
XGBoost	2.032
Support Vector Machine	2.184
ANN	2.413
Logistic Regression	2.635

5.4 CONCLUSION

The present chapter concludes the combined assessment of machine learning applications in groundwater quality prediction and soil contamination analysis around landfill affected conditions.

The groundwater modelling results demonstrated that machine learning models can effectively predict WA-WQI from selected hydro-chemical parameters. Among KNN, SVM, Decision Tree, Random Forest, XGBoost, and ANN, the Artificial Neural Network model achieved the best performance using all seven input features. The Taylor diagram also ranked ANN as the closest model to the reference point, followed by SVM and XGBoost. Validation confirmed the reliability of ANN as compared to other models.

These results confirm ANN as the most accurate model for complete feature-based WA-WQI prediction. The feature reduction analysis showed that groundwater quality prediction can also be performed using fewer input parameters. SVM maintained reliable performance with both four and two features. ANN using only Potassium also produced usable prediction results, although accuracy decreased compared with complete and reduced multi parameter cases. Hence, complete feature ANN remains the strongest model, while reduced feature SVM provides a cost-effective alternative for water quality monitoring where laboratory testing time and resource requirements need to be minimized.

The soil modelling results demonstrated that the Random Forest Genetic Algorithm model outperformed the comparison models for heavy metal contamination prediction. RF GA was compared with Random Forest, XGBoost, SVM, ANN, and Logistic Regression. The model also showed stronger performance across accuracy, recall, precision, sensitivity, specificity, False Negative Rate, and False Positive Rate when correlation informed Chi square feature selection was included. Accuracy improved from 92.978% for RF to 94.003% for RF GA with correlation features, while False Negative Rate decreased from 10.888% to 6.601%.

The convergence analysis confirmed the optimization strength of RF GA for soil contamination prediction. The RF GA model showed faster reduction in cost function values as compared to other models. This result demonstrates that Genetic Algorithm based optimization improved Random Forest performance by enhancing feature selection, reducing redundant inputs, and improving convergence behaviour.

Overall, the combined findings show that machine learning provides an effective approach for groundwater and soil contamination assessment in landfill affected environments. The combined framework reduces dependence on repeated laboratory testing, supports early identification of contamination risk, and provides a data driven basis for sustainable groundwater and soil management. Future research may include additional groundwater parameters such as heavy metals, nitrate, sulphate, aluminium, and phosphate, along with seasonal sampling and larger soil datasets from additional landfill affected locations.

CHAPTER 6: CONCLUSION, FUTURE SCOPE AND SOCIAL IMPACT

6.1 GENERAL

This chapter concludes the study on the assessment of landfill leachate impact on soil and groundwater quality around the Bhalswa landfill site by summarising the key findings and contributions of the research. It provides an overview of how the present study fulfils the research objectives and emphasizes its relevance for environmental monitoring, groundwater protection, landfill management. The chapter first recapitulates the major conclusions drawn from leachate characterization, soil and groundwater contamination assessment, pollution indices, SEM-EDX analysis, and machine learning-based prediction models. It then discusses the social impact of the study, particularly for landfill-adjacent communities, and finally outlines the future scope of research in the context of sustainable landfill management and environmental protection.

6.2 CONCLUSION

This section presents the major conclusions derived from the present research work. The key findings and overall outcomes of the study are summarized below.

- The characterization of Bhalswa landfill leachate confirmed its high pollution potential, as elevated physicochemical parameters and heavy metal concentrations indicate that leachate is a major source of contamination in the surrounding soil environment. The exceedance of Zn, Pb, Fe, Cd, and Mn concentrations above Indian Natural Background Soil limits, along with Cd exceeding the WHO soil limit, further confirms the impact of landfill-derived contamination on the soil quality.

- The present study identified major ions of Cd, Zn, and Pb for the geo-accumulation index, while the enrichment factor indicated an enrichment of Zn, Pb, Mn, and Cd, suggesting a strong landfill-related influence. The pollution indices clearly indicate metal contamination and anthropogenic enrichment in the soil affected by leachate. SEM-EDX analysis further supported these findings by providing elemental evidence of contamination within the soil matrix, confirming that leachate migration and waste decomposition have contributed to localized heavy metal accumulation.
- The overall geotechnical behaviour of the soil samples did not show severe or uniform deterioration across all boreholes. The soil profile remained predominantly sandy, and major changes were observed in local pockets having fine grained soils at deeper layers, where having higher moisture content, reduced dry density, lower permeability and lower shear strength. Thus, landfill leachate has significantly affected the soil chemistry and metal enrichment pattern, while its impact on geotechnical behaviour appears localized and comparatively moderate.
- Groundwater quality in the landfill-adjacent area showed severe degradation, indicating the possible influence of leachate migration on subsurface water resources. The heavy metal distribution pattern and Weighted Arithmetic Water Quality Index (WA-WQI) assessment collectively suggest that most groundwater samples were unsuitable for direct use without proper treatment, including for drinking, irrigation, and industrial purposes.
- Machine learning modelling proved effective for predicting groundwater quality using hydro-chemical parameters. Among the six developed models, ANN showed the best overall performance for complete feature-based WA-WQI prediction, while SVM also maintained reliable predictive ability with reduced input features. These results indicate that machine learning can support faster groundwater quality assessment and may reduce the need for extensive parameter monitoring in landfill-impacted areas.

- The hybrid RF+GA approach can be concluded as a reliable and practical tool for soil heavy metal prediction, as it combines the strength of Random Forest with the optimization capability of Genetic Algorithm. Its ability to achieve rapid convergence toward lower cost function values highlights its efficiency in feature selection and model optimization. Therefore, RF+GA may serve as a scalable and economical approach for forecasting soil pollution, supporting early contamination assessment, and assisting environmental management and policy formulation.

6.3 SOCIAL IMPACT

This study has significant social relevance as it addresses the scientific evidence that untreated landfill leachate can degrade nearby soil and groundwater quality through heavy metal enrichment and poor water quality conditions. This has direct implications for communities living near landfill sites, as contaminated groundwater and soil may affect drinking water safety, agricultural use, food-chain quality, and long-term human health. The study contributes to public health protection by identifying contamination levels and highlighting the need for treatment, monitoring and preventive management in landfill-adjacent areas. Although a detailed quantitative human health risk assessment was not carried out in the present work, the results provide an important scientific basis for future risk assessment involving exposure dose estimation, carcinogenic and non-carcinogenic risk evaluation, and ecological risk assessment. The soil and groundwater contamination data generated in this study can therefore support future studies based on established risk assessment frameworks, including WHO or USEPA based methodologies.

The findings directly align with SDG 3: Good Health and Well-being, as the assessment of heavy metals and degraded groundwater quality helps identify potential exposure risks for communities living near landfill sites. It also directly supports SDG 6: Clean Water and Sanitation by evaluating groundwater quality and highlighting the need for treatment before domestic, irrigation, or other

uses. Further, the study directly contributes to SDG 15: Life on Land by assessing soil contamination, heavy metal enrichment, and the possible impact of leachate on soil quality and terrestrial ecosystems. The contribution of research work to sustainable development goals (SDGs) are given in Table 6.1.

Table 6.1 contribution of research work to sustainable development goals (SDGs)

S. No.	SDG (Number / Official Title)	Relevance of the SDG to the Research Work – Chapter-wise / Research Publication-wise (as applicable)	Strength of Contribution
1.	**SDG 6** **Clean Water and Sanitation**	Chapters 3 and 5 / Publications: Bharadwaj et al. (2025a, 2025c, 2025d).	Strong
2.	**SDG 3** **Good Health and Well-Being**	Chapters 3, 4 and 6 / Publications: Bharadwaj et al. (2025a, 2025b).	Moderate
3.	**SDG 15** **Life on Land**	Chapters 3 and 5 / Publication: Bharadwaj et al. (2025b).	Strong
4.	**SDG 12** **Responsible Consumption and Production**	Chapters 3 and 6: soil and water (especially SDG Target 12.4). Publications: Bharadwaj et al. (2025a, 2025b).	Moderate
5.	**SDG 11** **Sustainable Cities and Communities**	Chapters 3, 4 and 6 / Publications: Bharadwaj et al. (2025a, 2025b).	Moderate

6.4 FUTURE SCOPE

1. Future studies may include long-term seasonal monitoring of landfill leachate, soil, and groundwater to understand temporal variation in contaminant migration during pre-monsoon, monsoon, and post-monsoon periods.
2. The study can be extended by increasing the number of sampling locations and depth intervals around the landfill to develop a more detailed spatial contamination profile and identify high-risk zones more accurately.
3. Future studies may include detailed human health risk assessment and ecological risk assessment for landfill-affected areas. Human health risk assessment may consider exposure pathways such as ingestion of contaminated groundwater, dermal contact, soil ingestion and food-chain transfer, along with risk quantification for carcinogenic and non-carcinogenic effects. Ecological risk assessment may also be carried out to evaluate the impact of heavy metal contamination on soil quality, vegetation, microorganisms and terrestrial ecosystems. Such assessments may be performed using established WHO or USEPA-based risk assessment frameworks.
4. Future work may also focus on developing and testing suitable leachate treatment and soil remediation strategies, such as adsorption, phytoremediation, bioremediation, or permeable reactive barriers, to reduce landfill-related contamination.
5. Future studies may evaluate the effectiveness of reactive barriers, geosynthetic liners, and soil amendments in controlling contaminant migration through poorly graded sandy soils.
6. Biological tracer studies, including spore-based tests and microbial indicators, may be conducted to better understand leachate movement pathways through permeable sandy soils toward groundwater.

REFERENCES

Abba, S. I., Yassin, M. A., Jibril, M. M., Tawabini, B., Soupios, P., Khogali, A., Shah, S. M. H., Usman, J., & Aljundi, I. H. (2024). Nitrate concentrations tracking from multi-aquifer groundwater vulnerability zones: Insight from machine learning and spatial mapping. Process Safety and Environmental Protection, 184, 1143–1157. https://doi.org/10.1016/j.psep.2024.02.041

Abd-Elmaboud, M. E., Saqr, A. M., El-Rawy, M., Al-Arifi, N., & Ezzeldin, R. (2024). Evaluation of groundwater potential using ANN-based mountain gazelle optimization: A framework to achieve SDGs in East El Oweinat, Egypt. Journal of Hydrology: Regional Studies, 52, 101703. https://doi.org/10.1016/j.ejrh.2024.101703

Adimalla, N., Qian, H., & Li, P. (2020). Entropy water quality index and probabilistic health risk assessment from geochemistry of groundwaters in hard rock terrain of Nanganur County, South India. Geochemistry, 80(4), 125544. https://doi.org/10.1016/j.chemer.2019.125544

Ahamad, A., Raju, N. J., Madhav, S., Gossel, W., & Wycisk, P. (2019). Impact of non-engineered Bhalswa landfill on groundwater from Quaternary alluvium in Yamuna flood plain and potential human health risk, New Delhi, India. Quaternary International, 507, 352–369. https://doi.org/10.1016/j.quaint.2018.06.011

Ahmed, S., Joshi, R., & Kumar, S. (2020). Seasonal variation of leachate quality at active landfill sites in Delhi, India. Proceedings of the Institution of Civil Engineers: Municipal Engineer, 173(3), 157–170. https://doi.org/10.1680/jmuen.17.00039

Alam, P., Sharholy, M., & Ahmad, K. (2020). A study on the landfill leachate and its impact on groundwater quality of Ghazipur Area, New Delhi, India. In A. S. Kalamdhad (Ed.), Recent

Developments in Waste Management: Lecture Notes in Civil Engineering 57, 345–357. Springer. https://doi.org/10.1007/978-981-15-0990-2_27.

Alghamdi, A. G., Aly, A. A., & Ibrahim, H. M. (2021). Assessing the environmental impacts of municipal solid waste landfill leachate on groundwater and soil contamination in western Saudi Arabia. Arabian Journal of Geosciences, 14(5). https://doi.org/10.1007/s12517-021-06583-9.

Anand, U., Reddy, B., Singh, V. K., Singh, A. K., Kesari, K. K., Tripathi, P., Kumar, P., Tripathi, V., & Simal-Gandara, J. (2021). Potential environmental and human health risks caused by antibiotic-resistant bacteria (ARB), antibiotic resistance genes (ARGs) and emerging contaminants (ECs) from municipal solid waste (MSW) landfill. Antibiotics, 10(4), 374. https://doi.org/10.3390/antibiotics10040374.

Angmo, S., & Shah, S. (2021). Impact of Okhla, Bhalswa and Ghazipur municipal waste dumpsites (landfill) on groundwater quality in Delhi. Current World Environment, 16(1). https://doi.org/10.12944/CWE.16.1.21.

APHA (American Public Health Association). (1990). Standard methods for the examination of water and wastewater (23rd ed.). Washington, DC: APHA.

Arafa, N. A., El-Said Salem, Z., Abdeldayem, A. L., Ghorab, M. A., Moustafa, Y. M., Soliman, S. A., Farag, M. H., Purohit, S., Elhag, M., & Youssef, Y. M. (2024). Advancing Deltaic Aquifer Vulnerability Mapping to Seawater Intrusion and Human Impacts in Eastern Nile Delta: Insights from Machine Learning and Hydrochemical Perspective. Earth Systems and Environment, 9(3), 2481–2506. https://doi.org/10.1007/s41748-024-00518-6

Awasthi, S.K. (2000) Prevention of Food Adulteration Act (Act No. 37 of 1954), Central and State Rules as Amended for 1999. 3rd Edition, Ashoka Law House, New Delhi.

Bai, M., et al. (2023). Removal of mixed contaminants from landfill leachate-contaminated soil by flushing with bio-surfactant: Laboratory column tests. Environmental Science and Pollution Research, 30(18), 53702–53711. https://doi.org/10.1007/s11356-023-26094-2

Balan, In., Madan Kumar, P., & Shivakumar, M. (2012). An assessment of groundwater quality using water quality index in Chennai, Tamil Nadu, India. Chronicles of Young Scientists, 3(2), 146. https://doi.org/10.4103/2229-5186.98688.

Bazoobandi, A., Emamgholizadeh, S., & Ghorbani, H. (2019). Estimating the amount of cadmium and lead in the polluted soil using artificial intelligence models. European Journal of Environmental and Civil Engineering, 26(3), 933–951. https://doi.org/10.1080/19648189.2019.1686429

Beinabaj, S. M. H., Heydariyan, H., Aleii, H. M., & Hosseinzadeh, A. (2023). Concentration of heavy metals in leachate, soil, and plants in Tehran's landfill: Investigation of the effect of landfill age on the intensity of pollution. Heliyon, 9(1), e13017. https://doi.org/10.1016/j.heliyon.2023.e13017

Bharadwaj, S., Gupta, A. K., & Sahu, A. K. (2025a). Characterization of Contaminated Soil and Groundwater Due to Landfill Leachate. Proceedings of the Indian Geotechnical Conference (IGC 2024), (Vol. 4, 91–10)7. https://doi.org/10.1007/978-981-96-7771-9_7

Bharadwaj, S., Gupta, A. K., & Sahu, A. K. (2025b). Effect of leachate contamination on soil. In A. Kashyap et al. (Eds.), Lecture Notes in Civil Engineering (Vol. 601, 98–108). Springer. https://doi.org/10.1007/978-981-96-4902-0_9

Bharadwaj, S., Gupta, A. K., & Sahu, A. K. (2025c). Tracing groundwater contaminants for sustainable development through machine learning algorithm. CLEAN – Soil, Air, Water, 53, e70083. https://doi.org/10.1002/clen.70083

Bharadwaj, S., Gupta, A. K., & Sahu, A. K. (2025d). Innovations in water quality management using machine learning approaches. Indian Journal of Engineering & Materials Sciences, 32(5), 536–551. https://doi.org/10.56042/ijems.v32i05.20867

Bhagat, S. K., Tiyasha, T., Awadh, S. M., Tung, T. M., Jawad, A. H., & Yaseen, Z. M. (2021). Prediction of sediment heavy metal at the Australian Bays using newly developed hybrid artificial intelligence models. Environmental Pollution, 268, 115663. https://doi.org/10.1016/j.envpol.2020.115663

Biswas, A. K., Kumar, S., Babu, S. S., Bhattacharyya, J. K., & Chakrabarti, T. (2010). Studies on environmental quality in and around municipal solid waste dumpsite. Resources, Conservation and Recycling, 55(2), 129–134. https://doi.org/10.1016/j.resconrec.2010.08.003

Bose, O., Das, P., Shaw, A., Layek, M. K., Smith, M., Sen, J., Sengupta, P., & Mukherjee, A. (2024). Land use and land cover patterns as a reflection of subsurface architecture groundwater quality in a large urban center (Varanasi) in the Ganges River Basin, India. Groundwater for Sustainable Development, 26, 101271. https://doi.org/10.1016/j.gsd.2023.101271.

Bouzayani, F., Aydi, A., & Abichou, T. (2014). Soil contamination by heavy metals in landfills: measurements from an unlined leachate storage basin. Environmental Monitoring and Assessment, 186(8), 5033–5040. https://doi.org/10.1007/s10661-014-3757-y

Buat-Ménard, P., & Chesselet, R. (1979). Variable influence of atmospheric flux on the trace metal chemistry of oceanic suspended matter. Earth and Planetary Science Letters, 42(3), 398–411. https://doi.org/10.1016/0012-821X(79)90049-9

Bureau of Indian Standards. (2012). Indian Standard IS 10500:2012 – Drinking water: Specification (2nd rev.). New Delhi, India: BIS. https://cpcb.nic.in/wqm/BIS_Drinking_Water_Specification.pdf

Bureau of Indian Standards. (1983a). Methods of sampling and test (physical and chemical) for water and wastewater (IS 3025 series). New Delhi, India: BIS.

Cao, W., & Zhang, C. (2020). Data prediction of soil heavy metal content by deep composite model. Journal of Soils and Sediments, 21(1), 487–498. https://doi.org/10.1007/s11368-020-02793-y

Central Ground Water Board. (2024). Ground water contamination studies around Bhalaswa landfill site, N.C.T., Delhi / Aquifer management plan: Bhalaswa landfill site, N.C.T. Delhi. Ministry of Jal Shakti, Government of India.

Che Nordin, N. F., Mohd, N. S., Koting, S., Ismail, Z., Sherif, M., & El-Shafie, A. (2021). Groundwater quality forecasting modelling using artificial intelligence: A review. In Groundwater for Sustainable Development 14 (100643). Elsevier BV. https://doi.org/10.1016/j.gsd.2021.100643

Drall, J. K., Rautela, R., Jambhulkar, R., Kataria, A. K., & Kumar, S. (2025). Effect of heavy metals contamination due to leachate migration from uncontrolled dumpsites: A comprehensive analysis on soil and groundwater. Journal of Environmental Management, 373, 123473. https://doi.org/10.1016/j.jenvman.2024.123473

El-Said Salem, Z., Arafa, N. A., Abdeldayem, A. L., & Youssef, Y. M. (2025). Machine learning-enhanced GALDIT modeling for the Nile Delta aquifer vulnerability assessment in the Mediterranean region. Groundwater for Sustainable Development, 28, 101403. https://doi.org/10.1016/j.gsd.2024.101403

Falowo, O. O., Otuaga, M. P., & Lasore, E. A. (2023). Groundwater prediction analysis for sustainable development in Okeigbo, Southwestern Nigeria: A GIS supported AHP approach. Indonesian Journal of Earth Sciences, 3(2), A684.

Gaur, V. K., Gautam, K., Vishvakarma, R., Sharma, P., Pandey, U., Srivastava, J. K., Varjani, S., Chang, J.-S., Ngo, H. H., & Wong, J. W. C. (2024). Integrating advanced techniques and machine learning for landfill leachate treatment: Addressing limitations and environmental concerns. Environmental Pollution, 354, 124134. https://doi.org/10.1016/j.envpol.2024.124134

GodvinSharmila, V., Shanmugavel, S. P., Tyagi, V. K., & Rajesh Banu, J. (2023). Microplastics as emergent contaminants in landfill leachate: Source, potential impact and remediation technologies. Journal of Environmental Management, 343, 118240. https://doi.org/10.1016/j.jenvman.2023.118240

Gowd, S. S., Reddy, M. R., & Govil, P. K. (2010). Assessment of heavy metal contamination in soils at Jajmau (Kanpur) and Unnao industrial areas of the Ganga Plain, Uttar Pradesh, India. Journal of Hazardous Materials, 174(1–3), 113–121. https://doi.org/10.1016/j.jhazmat.2009.09.024

Haghiabi, A. H., Nasrolahi, A. H., & Parsaie, A. (2018). Water quality prediction using machine learning methods. Water Quality Research Journal, 53(1), 3–13. https://doi.org/10.2166/wqrj.2018.025

Harun, S. N., Ali Rahman, Z., & Damanhuri, N. A. (2025). Impact of landfill leachate contamination on the geotechnical behavior of granitic sandy-clayey soil. International Journal of Environmental Science and Technology, 22, 12423–12436. https://doi.org/10.1007/s13762-025-06622-y

Harun, S. N., Ali Rahman, Z., Rahim, S. A., Lihan, T., & Idris, W. M. R. (2013). Effects of leachate on geotechnical characteristics of sandy clay soil. AIP Conference Proceedings, 1571, 530–536. https://doi.org/10.1063/1.4858709

Hussein, M., Yoneda, K., Mohd-Zaki, Z., Amir, A., & Othman, N. (2021). Heavy metals in leachate, impacted soils and natural soils of different landfills in Malaysia: An alarming threat. Chemosphere, 267, 128874. https://doi.org/10.1016/j.chemosphere.2020.128874

Islam, S., Bano, H., Malik, A. A., & Alotaibi, F. (2024). Landfill leachate: An invisible threat to soil quality of temperate Himalayas. PLOS ONE, 19(11), e0314006. https://doi.org/10.1371/journal.pone.0314006

Jamshidzadeh, Z., Ehteram, M., & Shabanian, H. (2024). Bidirectional Long Short-Term Memory (BILSTM) - Support Vector Machine: A new machine learning model for predicting water quality parameters. Ain Shams Engineering Journal, 15(3), 102510. https://doi.org/10.1016/j.asej.2023.102510.

Javadi, F., Qaderi, K., Ahmadi, M. M., Rahimpour, M., Madadi, M. R., & Mahdavi-Meymand, A. (2022). Application of classical and novel integrated machine learning models to predict sediment discharge during free-flow flushing. In Scientific Reports (Vol. 12, Issue 1). Springer Science and Business Media LLC. https://doi.org/10.1038/s41598-022-23781-x

Jia, X., Hu, B., Marchant, B. P., Zhou, L., Shi, Z., & Zhu, Y. (2019). A methodological framework for identifying potential sources of soil heavy metal pollution based on machine learning: A case study in the Yangtze Delta, China. Environmental Pollution, 250, 601–609. https://doi.org/10.1016/j.envpol.2019.04.047

Kanmani, S., & Gandhimathi, R. (2013). Assessment of heavy metal contamination in soil due to leachate migration from an open dumping site. Applied Water Science, 3(1), 193–205. https://doi.org/10.1007/s13201-012-0072-z

Krishan, A., Mishra, R. K., & Khursheed, A. (2022). Assessment of water quality using water quality index: A case study of the River Gomti, Lucknow, Uttar Pradesh, India. Urban Water Journal, 19(5), 520–530. https://doi.org/10.1080/1573062x.2022.2032210

Keçeci, M., Gökmen, F., Usul, M., Koca, C., & Uygur, V. (2024). Prediction of cadmium content using machine learning methods. Environmental Earth Sciences, 83, 362. https://doi.org/10.1007/s12665-024-11672-5

Kenue, A. W., Sana, E., Adhikary, A., Pal, S., & Hussain, M. (2023). Assessment of landfill leachate and its surface and subsurface migration: Experimental and numerical investigations. Journal of Hazardous, Toxic, and Radioactive Waste, 27(4), 05023003. https://doi.org/10.1061/JHTRBP.HZENG-1227

Khodary, S. M., Negm, A. M., & Tawfik, A. (2018). Geotechnical properties of the soils contaminated with oils, landfill leachate, and fertilizers. Arabian Journal of Geosciences, 11, 13. https://doi.org/10.1007/s12517-017-3372-7

Koda, E., Miszkowska, A., Sieczka, A., & Osiński, P. (2020). Heavy metals contamination within restored landfill site in Poland. Environmental Geotechnics, 7(7), 512–521. https://doi.org/10.1680/jenge.18.00031

Kooch, Y., Nouraei, A., Haghverdi, K., Kolb, S., & Francaviglia, R. (2023). Landfill leachate has multiple negative impacts on soil health indicators in Hyrcanian forest, northern Iran. Science of the Total Environment, 896, 166341. https://doi.org/10.1016/j.scitotenv.2023.166341

Kooch, Y., Nouraei, A., Wang, L., Wang, X., Wu, D., Francaviglia, R., Frouz, J., & Parsapour, M. K. (2024). Long-term landfill leachate pollution suppresses soil health indicators in natural ecosystems of a semi-arid environment. Chemosphere, 367, 143647. https://doi.org/10.1016/j.chemosphere.2024.143647

Kumar, D., & Alappat, B. J. (2003). Development of a leachate pollution index for municipal solid waste landfills. Waste Management, 23(7), 557–561.

Kumar, D., & Alappat, B. J. (2005a). Analysis of Leachate Pollution Index and formulation of sub-leachate pollution indices. Waste Management & Research, 23(3), 230–239. https://doi.org/10.1177/0734242X05054875

Kumar, D., & Alappat, B. J. (2005b). Evaluating leachate contamination potential of landfill sites using Leachate Pollution Index. Clean Technologies and Environmental Policy, 7, 190–197. https://doi.org/10.1007/s10098-004-0269-4

Kumar, D., & Alappat, B. J. (2005c). Analysis of Leachate Pollution Index and formulation of sub-leachate pollution indices. Waste Management & Research, 23(3), 230–239. https://doi.org/10.1177/0734242X05054875

Kumari, M., & Rai, S. C. (2020). Hydrogeochemical Evaluation of Groundwater Quality for Drinking and Irrigation Purposes Using Water Quality Index in Semi Arid Region of India. Journal of the Geological Society of India, 95(2), 159–168. https://doi.org/10.1007/s12594-020-1405-4

Kuhad, M. S., Malik, R. S., Singh, A., & Dahiya, I. S. (1989). Background levels of heavy metals in agricultural soils of Indo-Gangetic Plains of Haryana. Journal of the Indian Society of Soil Science, 37, 700–705.

Li, D., Zhang, Y., Yu, F., Wang, J., Zhang, X., Feng, L., Lang, T., & Yang, F. (2024). Vadose-zone characteristic pollutants distribution, microbial community structure and functionality changes in response to long-term leachate pollution of an informal landfill site. Science of the Total Environment, 948, 174596. https://doi.org/10.1016/j.scitotenv.2024.174596

Li, P., Karunanidhi, D., Subramani, T., & Srinivasamoorthy, K. (2021). Sources and consequences of groundwater contamination. Archives of Environmental Contamination and Toxicology, 80(1), 1–10. https://doi.org/10.1007/s00244-020-00805-z

Lott, D., Darioosh, R. P., Weiksnar, K., Laux, S., & Townsend, T. G. (2025). A comparison of bulk inorganic constituents and trace pollutant concentration in leachates by landfill type. Waste Management & Research, 43(8), 1321–1331. https://doi.org/10.1177/0734242X251322146

Lu, H., & Ma, X. (2020). Hybrid decision tree-based machine learning models for short-term water quality prediction. Chemosphere, 249, 126169. https://doi.org/10.1016/j.chemosphere.2020.126169

Maghrebi, M., Noori, R., Partani, S., Araghi, A., Barati, R., Farnoush, H., & Torabi Haghighi, A. (2021). Iran's Groundwater Hydrochemistry. Earth and Space Science, 8(8). Portico. https://doi.org/10.1029/2021ea001793

Mir, B. A., Wani, A. R., & Wani, K. M. N. S. (2025). Influence of municipal solid waste leachate on the geotechnical properties of soil. Indian Geotechnical Journal, 56, 1403–1422. https://doi.org/10.1007/s40098-025-01175-w

Ministry of Environment, Forest and Climate Change. (2016). Solid Waste Management Rules, 2016. Government of India.

Ministry of Water Resources. (2011). National Water Mission under the National Action Plan on Climate Change: Comprehensive mission document (Vol. 1). Government of India.

Müller, G. (1969). Index of geoaccumulation in sediments of the Rhine River. GeoJournal, 2(3), 108–118.

Nayan, A.-A., Kibria, M. G., Rahman, Md. O., & Saha, J. (2020). River Water Quality Analysis and Prediction Using GBM. 2020 2nd International Conference on Advanced Information and Communication Technology (ICAICT), 219–224. https://doi.org/10.1109/icaict51780.2020.9333492

Niazkar, M., Menapace, A., Brentan, B., Piraei, R., Jimenez, D., Dhawan, P., & Righetti, M. (2024). Applications of XGBoost in water resources engineering: A systematic literature review (Dec 2018–May 2023). In Environmental Modelling & Software (Vol. 174, p. 105971). Elsevier BV. https://doi.org/10.1016/j.envsoft.2024.105971

Paes, É. de C., Veloso, G. V., Fonseca, A. A. da, Fernandes-Filho, E. I., Fontes, M. P. F., & Soares, E. M. B. (2022). Predictive modeling of contents of potentially toxic elements using morphometric data, proximal sensing, and chemical and physical properties of soils under mining influence. Science of the Total Environment, 817, 152972. https://doi.org/10.1016/j.scitotenv.2022.152972

Parakh, D. (2022). Groundwater regulation bills in Haryana—A call for groundwater conservation and management for sustainable irrigation: Opportunities and challenges. Journal of Coastal Life Medicine, 10(2), 642–651.

Qian, Y., Hu, P., Lang-Yona, N., Xu, M., Guo, C., & Gu, J. D. (2024). Global landfill leachate characteristics: Occurrences and abundances of environmental contaminants and the microbiome. Journal of Hazardous Materials, 461, 132446. https://doi.org/10.1016/j.jhazmat.2023.132446

Qiao, F., Wang, J., Chen, Z., Zheng, S., Kwaw, A. K., Zhao, Y., & Huang, J. (2025). Organic contamination pressure shapes spatiotemporal variability of shallow groundwater bacterial communities and temporal patterns when facing new environmental disturbances. Journal of Hydrology, 653, 132764. https://doi.org/10.1016/j.jhydrol.2025.132764

Raheja, H., Goel, A., & Pal, M. (2022). An evaluation of groundwater quality and its suitability for drinking and irrigation uses under the GIS framework. Water Practice & Technology, 17(11), 2259–2277. https://doi.org/10.2166/wpt.2022.134

Rani, A., Negi, S., Hussain, A., & Kumar, S. (2020). Treatment of urban municipal landfill leachate utilizing garbage enzyme. Bioresource Technology, 297, 122437. https://doi.org/10.1016/j.biortech.2019.122437

Robey, N. M., Liu, Y., Crespo-Medina, M., Bowden, J. A., Solo-Gabriele, H. M., Townsend, T. G., & Tolaymat, T. (2024). Characterization of per- and polyfluoroalkyl substances (PFAS) and other constituents in MSW landfill leachate from Puerto Rico. Chemosphere, 358, 142141. https://doi.org/10.1016/j.chemosphere.2024.142141

Saghi, M. H., Nadimi, H., Eslami, A., Alavi Bakhtiarvand, S. N., Oghazyan, A., Setoudeh, S., & Sargolzaei, M. S. (2024). Characteristics and pollution indices of leachates from municipal solid waste landfills in Iranian metropolises and their implications for MSW management. Scientific Reports, 14, 27285. https://doi.org/10.1038/s41598-024-78630-w

Saha, P., Saikia, K. K., Kumar, M., & Handique, S. (2023). Assessment of health risk and pollution load for heavy and toxic metal contamination from leachate in soil and groundwater in the vicinity of dumping site in Mid-Brahmaputra Valley, India. Total Environment Research Themes, 8, 100076. https://doi.org/10.1016/j.totert.2023.100076

Sathiyamurthi, S., Subbarayan, S., Ramya, M., Sivasakthi, M., Gobi, R., Qaysi, S., Praveen Kumar, S., Lee, J., Alarifi, N., Wahba, M., & M. Youssef, Y. (2024). Advancing Agricultural Land Suitability in Urbanized Semi-Arid Environments: Insights from Geospatial and Machine Learning Approaches. ISPRS International Journal of Geo-Information, 13(12), 436. https://doi.org/10.3390/ijgi13120436

Sharma, A., Ganguly, R., & Gupta, A. K. (2020). Impact assessment of leachate pollution potential on groundwater: An indexing method. Journal of Environmental Engineering, 146(3), 05019007. https://doi.org/10.1061/(ASCE)EE.1943-7870.0001647

Sharma, M., Khare, M., & Mishra, R. K. (2024). Air quality changes in Delhi due to open waste burning: An accidental fire in Bhalswa landfill. International Journal of Environmental Science and Technology, 21, 655–664. https://doi.org/10.1007/s13762-023-04921-w

Sidhu, B. S., Sharma, D., Tuteja, T., Gupta, S., & Kumar, A. (2015). Human health risk assessment of heavy metals from Bhalaswa landfill, New Delhi, India. In N. J. Raju et al. (Eds.), Management of Natural Resources in a Changing Environment (pp. 215–223). Springer. https://doi.org/10.1007/978-3-319-12559-6_16

Singh, V., Pencina, M., Einstein, A.J., Liang, J.X., Berman, D.S., Slomka, P.: Impact of train/test sample regimen on performance estimate stability of machine learning in cardiovascular imaging. Sci. Rep. 11, 14490 (2021). https://doi.org/10.1038/s41598-021-93651-5

Song, N., Zhang, Q., Wang, Y., Gao, L., Liu, S., Yao, H., Liu, R., & Xu, H. (2023). Investigation on molecular characteristics of organic compounds during a full scale landfill leachate treatment process based on non targeted analysis. Environmental Research, 238, 117258. https://doi.org/10.1016/j.envres.2023.117258

Srisuradetchai, P., & Suksrikran, K. (2024). Random kernel k-nearest neighbors regression. Frontiers in Big Data, 7. https://doi.org/10.3389/fdata.2024.1402384

Srivastava, S. K., & Ramanathan, A. L. (2008). Geochemical assessment of groundwater quality in vicinity of Bhalswa landfill, Delhi, India, using graphical and multivariate statistical methods. Environmental Geology, 53(7), 1509–1528. https://doi.org/10.1007/s00254-007-0762-2

Sutherland, R. A., Tolosa, C. A., Tack, F. M. G., & Verloo, M. G. (2000). Characterization of selected element concentrations and enrichment ratios in background and anthropogenically impacted roadside areas. Archives of Environmental Contamination and Toxicology, 38, 428–438. https://doi.org/10.1007/s002440010057

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. In Journal of Geophysical Research: Atmospheres (Vol. 106, Issue D7, pp. 7183–7192). American Geophysical Union (AGU). https://doi.org/10.1029/2000jd900719

United Nations. (2018). Sustainable Development Goal 6 synthesis report 2018 on water and sanitation. https://doi.org/10.18356/e8fc060b-en

United Nations. (2015). Transforming our world: The 2030 agenda for sustainable development. United Nations.

United Nations Educational, Scientific and Cultural Organization. (2024). The United Nations world water development report 2024: Water for prosperity and peace.

UNESCO World Water Assessment Programme. (2022). The United Nations world water development report 2022: Groundwater: Making the invisible visible. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000380726

Vasanthavigar, M., Srinivasamoorthy, K., Vijayaragavan, K., Rajiv Ganthi, R., Chidambaram, S., Anandhan, P., Manivannan, R., & Vasudevan, S. (2010). Application of water quality index for groundwater quality assessment: Thirumanimuttar sub-basin, Tamilnadu, India. Environmental Monitoring and Assessment, 171(1–4), 595–609. https://doi.org/10.1007/s10661-009-1302-1

Wang, H., Yilihamu, Q., Yuan, M., Bai, H., Xu, H., & Wu, J. (2020). Prediction models of soil heavy metal(loid)s concentration for agricultural land in Dongli: A comparison of regression and random forest. Ecological Indicators, 119, 106801. https://doi.org/10.1016/j.ecolind.2020.106801

Wisitthammasri, W., Promduang, P., & Chotpantarat, S. (2024). Characterization of microplastics in soil, leachate and groundwater at a municipal landfill in Rayong Province, Thailand. *Journal of Contaminant Hydrology, 267*, 104455. https://doi.org/10.1016/j.jconhyd.2024.104455

WHO/FAO (World Health Organization/Food and Agriculture Organization) (2007) Joint FAO/WHO Food Standards Programme Codex Alimentarius Commission Twenty-Seventh Session Report of the Thirty-Sixth Session of the Codex Committee on Food Hygiene.

World Meteorological Organization. (2022, March 22). World Water Day focuses on groundwater.

Yang, H., Huang, K., Zhang, K., Weng, Q., Zhang, H., & Wang, F. (2021). Predicting heavy metal adsorption on soil with machine learning and mapping global distribution of soil adsorption capacities. Environmental Science & Technology, 55(20), 14316–14328. https://doi.org/10.1021/acs.est.1c02479

Yang, L., Shami, A.: On hyperparameter optimization of machine learning algorithms: Theory and practice. Neurocomputing. 415, 295–316 (2020) https://doi.org/10.1016/j.neucom.2020.07.061

Yoshida, M., Ahmed, S., Nebil, S., & Ahmed, G. (2002). Characterization of leachate from Henchir El Yahoudia close landfill. Water Waste Environ Res, 1(2), 129-42.

Yousefi Kebria, D., Ghavami, M., Javadi, S., & Goharimanesh, M. (2018). Combining an experimental study and ANFIS modeling to predict landfill leachate transport in underlying soil: A case study in North Iran. Environmental Monitoring and Assessment, 190, 26. https://doi.org/10.1007/s10661-017-6374-8

Yu, X., Lyu, S., Zhao, W., Guo, C., Xu, J., & Sui, Q. (2024). A picture of pharmaceutical pollution in landfill leachates: Occurrence, regional differences and influencing factors. Waste Management, 184, 20–27. https://doi.org/10.1016/j.wasman.2024.05.019

Zahedi, S. (2017). Modification of expected conflicts between Drinking Water Quality Index and Irrigation Water Quality Index in water quality ranking of shared extraction wells using Multi Criteria Decision Making techniques. Ecological Indicators, 83, 368–379. https://doi.org/10.1016/j.ecolind.2017.08.017

Zairi, M., Ferchichi, M., Ismaı¨l, A., Jenayeh, M., & Hammami, H. (2004). Rehabilitation of El Yahoudia dumping site, Tunisia. Waste Management, 24(10), 1023–1034. https://doi.org/10.1016/j.wasman.2004.07.002

Zhang, H., Yin, S., Chen, Y., Shao, S., Wu, J., Fan, M., Chen, F., & Gao, C. (2020). Machine learning-based source identification and spatial prediction of heavy metals in soil in a rapid

urbanization area, eastern China. Journal of Cleaner Production, 273, 122858. https://doi.org/10.1016/j.jclepro.2020.122858

Zhang, Y., Lei, M., Li, K., & Ju, T. (2023). Spatial prediction of soil contamination based on machine learning: A review. Frontiers of Environmental Science & Engineering, 17(8). https://doi.org/10.1007/s11783-023-1693-1

CLEAN – Soil, Air, Water

WILEY-VCH

CLEAN Soil Air Water

RESEARCH ARTICLE

Tracing Groundwater Contaminants for Sustainable Development Through Machine Learning Algorithm

Shivangi Bharadwaj | Ashok Kumar Gupta | Anil Kumar Sahu

Department of Civil Engineering, Delhi Technological University, New Delhi, India

Correspondence: Shivangi Bharadwaj (shivangi.bharadwaj23@gmail.com)

Received: 30 April 2025 | **Revised:** 27 November 2025 | **Accepted:** 8 December 2025

Keywords: artificial neural network (ANN) | groundwater monitoring | hyperparameter optimization | machine learning techniques | water quality index

ABSTRACT

Groundwater (GW) contamination and its monitoring are continuously posing a significant challenge in cities like Delhi and its surroundings territories. The study analyzed seven hydro-chemical parameters such as calcium, chloride, magnesium, pH, potassium, total hardness (TH), and electrical conductivity (EC) for 488 GW samples to calculate Weighted Arithmetic Water Quality Index (WA-WQI), and further prediction of water quality was performed using the machine learning (ML) approaches. This study employed six ML models, artificial neural network (ANN), random forest (RF), *K*-nearest neighbors (KNN), decision tree (DT), extreme gradient boosting (XGBoost), and support vector machine (SVM) to predict WA-WQI. The prediction performance was assessed using R^2, root mean square error (RMSE), mean absolute error (MAE), centered RMSE (CRMSE), and mean absolute percentage error (MAPE). The obtained results revealed that ANN demonstrated the highest accuracy with lowest errors in MAE of 4.518, MAPE of 3.276, RMSE of 8.294, and CRMSE of 8.236. The ANN model exhibited the highest R^2 value of 0.99 among all examined models. The performance evaluation order found was ANN > SVM > XGBoost > RF > KNN > DT, and Taylor's diagram was employed to evaluate the efficacy of each model. The validation results show that ANN models with specific samples yielded promising results with R^2 value of 0.99, MAE of 3.746, MAPE of 2.550, RMSE of 5.520, and CRMSE of 5.373. The outcomes of the study indicate that ANN is the best performing model, providing a cost-efficient alternative for predicting GW quality, significantly reducing extensive laboratory testing, offering practical support for real-time monitoring, and helping policy makers to take decisions for mitigation and improving water quality.

1 | Introduction

Water bodies cover around 71% of the Earth's surface, making up to a total area of 1386 million km^3, of which only 0.5% is accessible freshwater [1]. Water scarcity is a serious problem; despite its abundance, approximately half the world's population will face severe shortages at least part of the year [2]. The global need for water probably rises by 20%–30% by 2050 due to population growth and economic expansion, increasing the strain on these finite resources [3]. Groundwater (GW) provides more freshwater than lakes and streams but is hard to collect, slow to recover, and easily contaminated. GW supplies 25% of irrigation water and half of residential freshwater [2]. The composition of the water plays an important role in determining the quality of water for consumption purpose [4]. The physical and chemical properties of GW have degraded due to its overexploitation in regions with high levels of climate variability.

Globally, GW contamination is a significant issue that has major impact on public health [5]. Delhi and the National Capital Region (NCR) of India, encompassing portions of Haryana, Rajasthan, and Uttar Pradesh, experience considerable pressure

© 2025 Wiley-VCH GmbH

Indian Journal of Engineering & Materials Sciences
Vol. 32, October 2025, pp. 536-551
DOI: 10.56042/ijems.v32i05.20867

Innovations in water quality management using machine learning approaches

Shivangi Bharadwaj*, Ashok Kumar Gupta, & Anil Kumar Sahu
Department of Civil Engineering, Delhi Technological University, New Delhi 110 042, India

Received: 21 June 2025; accepted: 29 August 2025

Groundwater contamination has been posing a significant threat to sustainable water resource management, particularly in industrialized and urbanized regions. This research has introduced a novel, data-driven framework that integrates machine learning, statistical data analysis, and feature optimization to evaluate and forecast groundwater quality. Analytical results of 488 groundwater samples had been tested, and four feature reduction scenarios had been implemented using Pearson correlation to evaluate predictive performance with minimal input variables. Statistical analysis has highlighted elevated levels of parameters such as Electrical Conductivity, Chloride, Magnesium, and Total Hardness, exceeding permissible limits, and have been causing most samples to be unsuitable for consumption without treatment. To enhance groundwater monitoring and reduce laboratory testing costs, six machine learning algorithms, K-Nearest Neighbors, Support Vector Machine, Decision Tree, Random Forest, XGBoost, and Artificial Neural Network, have been used to predict the Weighted Arithmetic Water Quality Index. Model accuracy had been tested using statistical metrics such as R^2, RMSE, MAE, MAPE, and CRMSE, with effectiveness assessed using Taylor diagrams. ANN exhibited the highest accuracy even when using a single input (K), while SVM maintained consistent reliability with only two inputs (Mg and K), providing a cost-effective monitoring solution. Validation with 70 independent datasets has confirmed the robustness and applicability of the suggested methodology. The study has presented an innovative modeling strategy that has substantially decreased laboratory testing needs while preserving predictive reliability. Additionally, it has offered practical implications for scalable, cost-effective deployment in areas with water scarcity or insufficient dataset.

Keywords: Groundwater contamination, K-fold cross validation, Machine learning models, Regional hydrology, Water quality index

1 Introduction

Groundwater constitutes approximately 98% of the global freshwater reserves, which are extensively used for household, agricultural, and industrial purposes. The chemical composition of groundwater is crucial for evaluating its suitability for human consumption[1-2]. Geological characteristics and environmental factors affect the quality of groundwater. Globally, groundwater contamination is a significant issue that presents grave risks to public health and environmental quality[3]. A Water quality index (WQI) is a quantitative measure used to assess water quality for various purposes. It can be used to evaluate the appropriateness of water for drinking, industrial uses, aquatic life, and other applications. Greater WQI values indicate superior water quality[4]. Water Quality Classification (WQC) is a system that classifies water as either mildly contaminated or pure based on the WQI value range[5]. Agriculture serves as the primary consumer of water resources[6-7]. The over use of groundwater in areas with significant climate variability has led to the deterioration of this resource in terms of both physical and chemical properties.

Purified water is an essential resource upon which living creatures depend. Consequently, the development of a water quality forecasting technique to predict future water quality conditions has great social and economic importance. Predictive models allow for real-time or near-real-time estimation of the WQI and WQC, making them more efficient and cost-effective than traditional laboratory analyses. This capability enables continuous monitoring of water quality, early identification of declines, and prompt responses to potential hazards or pollution events.

A number of researchers have assessed the potability of groundwater using the following major parameters, pH, Total Hardness (TH), Electrical Conductivity (EC), Magnesium (Mg^{2+}), calcium (Ca^{2+}), sodium (Na^{+}), potassium (K^{+}), bicarbonate (HCO_3^{-}), nitrate (NO_3^{2-}), fluoride (F), and sulphate (SO_4^{2-})[8-11]. Because measuring any of these factors in groundwater is a time-consuming and costly process,

*Corresponding author
(Email: shivangibharadwaj_2k21phdce03@dtu.ac.in)

Check for updates

CLEAN – Soil, Air, Water

WILEY-VCH

CLEAN Soil Air Water

RESEARCH ARTICLE

Enhancement of Landfill Leachate Attenuation in Sandy Soil Using Compost Amendment: A Soil Column Study

Jyotirmay | Arin Yadav | Jyoti Meena | Shivangi Bharadwaj | Ashok Kumar Gupta

Department of Civil Engineering, Delhi Technological University, New Delhi, India

Correspondence: Shivangi Bharadwaj (shivangi.bharadwaj23@gmail.com)

Received: 11 March 2026 | **Revised:** 4 June 2026 | **Accepted:** 11 June 2026

Keywords: compost amendment | controlled column study | heavy metals | landfill leachate | natural attenuation

ABSTRACT

Soil's attenuation capacity is essential for protecting structures and ecosystems from natural and anthropogenic disturbances. This study investigates the natural attenuation capacity of soil using controlled laboratory column experiments to demonstrate its ability to adsorb, retain, and partially degrade leachate contaminants, thereby limiting their transport to groundwater resources. This study used two vertical soil columns: one containing poorly graded sand and the other containing the same soil enhanced with a 3:2 compost-to-soil ratio to increase attenuation. The leachate samples show high levels of organic pollutants, inorganic constituents, and heavy metals. The preliminary metal concentration pattern in the initial soil followed the order Fe > Al > Zn > Cr > Cu > Pb > Ni ≈ Cd, indicating the dominance of Fe and Al among native soil constituents. The pretreated leachate was percolated through both columns for 30 days under ambient laboratory conditions. The influent and effluent samples from the columns were tested for pH, electrical conductivity (EC), ammoniacal nitrogen, sulfate, phosphate, nitrate, biochemical oxygen demand (BOD), chemical oxygen demand (COD), total suspended solids (TSS), total dissolved solids (TDS), alkalinity, total hardness (TH), calcium, chloride, and heavy metals. The results reveal that the compost-amended soil column removed more heavy metals than the natural column. The soil medium's capacity to retain nutrients is demonstrated by the consistent decrease in phosphorus concentration from leachate to effluent. Soil modification altered pH, electrical conductivity, and hardness. BOD, COD, and chloride concentrations dropped significantly in the modified column. Results indicate that natural soil possesses a significant ability to reduce leachate pollutants, but adding organic amendments, like compost, boosts its contaminant-removal efficiency.

1 | Introduction

Municipal solid waste (MSW) generation, driven by rapid urbanization, population growth, and changing consumption patterns, has emerged as a critical environmental challenge worldwide [1, 2]. In India, landfilling remains the most prevalent MSW disposal method because of its economic feasibility. However, many landfill sites operate without engineered liners, leachate collection systems, or adequate treatment facilities [3]. Due to this, landfill leachate formed by the percolation of precipitation and moisture through decomposing waste poses a significant threat to surrounding soil and groundwater systems [4]. Landfill leachate is characterized by high concentrations of dissolved salts, ammoniacal nitrogen, organic matter, and potentially toxic

Abbreviations: Al, aluminum; BOD, biochemical oxygen demand; Cd, cadmium; CEC, cation exchange capacity; COD, chemical oxygen demand; Cr, chromium; Cu, copper; EC, electrical conductivity; Fe, iron; IARC, International Agency for Research on Cancer; MSW, municipal solid waste; Ni, nickel; Pb, lead; TA, total alkalinity; TDS, total dissolved solids; TH, total hardness; TOC, total organic carbon; TOM, total organic matter; TSS, total suspended solids; Zn, zinc.

© 2026 Wiley-VCH GmbH

Z. Dt. Ges. Geowiss. (J. Appl. Reg. Geol.)
Published online August 2026

Early Access Article

Optimised machine learning model for predicting heavy metal contamination in soil around a landfill

Shivangi Bharadwaj[1], Ashok Kumar Gupta[2] & Anil Kumar Sahu[3]*

Bharadwaj, S., Gupta, A. K. & Sahu, A. K. (2026): Optimised machine learning model for predicting heavy metal contamination in soil around a landfill. – Z. Dt. Ges. Geowiss.

Abstract: This study investigates the environmental and geotechnical implications of leachate migration from the Bhalswa landfill in New Delhi, India, which has accumulated over 8 million tons of mixed, unsegregated waste over the decades. Potential contamination risks to the surrounding soil were assessed by collecting 150 soil samples from ten strategically positioned boreholes around the landfill, at depths of 1 to 15 metres. Laboratory analyses were conducted to determine physicochemical properties and heavy metal concentrations. The results revealed significant vertical and horizontal variations in parameters like pH, EC, TDS, chlorides, alkalinity, total hardness, and concentrations of Mg, Al, Mn, Fe, Zn, Cu, Cd, Cr, and Pb. This research objective is to contribute to computational geomechanics by combining site-specific subsurface characterisation with advanced machine learning to assess the geotechnical impacts of contaminant transport. Elevated levels were particularly evident at deeper layers and nearer to the landfill, indicating active leachate percolation into subsurface zones. Due to the dataset's multidimensionality, traditional predictive methods proved insufficient in terms of scalability and robustness. A hybrid machine learning model combining Random Forest (RF) and Genetic Algorithm (GA) was created to address these gaps. The RF-GA model was benchmarked against five established algorithms – RF, XGBoost, Support Vector Machines (SVM), Artificial Neural Networks (ANN), and Logistic Regression (LR). The RF-GA model consistently outperformed the others, particularly with larger training datasets, demonstrating high specificity and rapid convergence to lower cost function values. This research highlights the critical role of machine learning in environmental risk assessment and offers a scalable, data-driven solution for sustainable waste management practices. By incorporating advanced feature selection and optimisation techniques, this study sets a precedent for leveraging machine learning in geotechnical engineering and offers innovative solutions for environmental risk mitigation.

Keywords: RF-GA hybrid model, heavy metals contamination, soil behaviour, landfill leachate, contaminant transport

1. Introduction

Open dumping is the predominant practice for disposing of solid waste in many underdeveloped countries. These unscientific landfills, often referred to as open dumps, are typically constructed without liner systems and lack proper measures to manage water flow (Bharadwaj et al. 2025a). Consequently, they lead to significant environmental challenges, including the contamination of soil, surface water, and groundwater (Ahmed et al. 2020). Soil pollution from landfills disrupts ecosystems and poses risks to human health and the well-being of nearby flora and fauna (Kanmani & Gandhimathi 2013). While landfills are essential for waste management, they inadvertently become sources of pollution due to the decomposition and chemical reactions of discarded materials. Organic waste decomposition produces harmful byproducts such as methane gas and leachate, a toxic liquid formed when water percolates through waste. Leachate contains hazardous substances, including toxic metals, pesticides, and other chemicals, which infiltrate the soil, causing progressive pollution (Alghamdi et al. 2021). Open dumps are unattractive, unclean, and stinky. They attract pests, including rodents, insects, pigs, and scavengers. Percolating water can carry harmful chemicals from dumpsites, exacerbating pollution issues (Raman et al. 2008). Municipal solid waste, coupled with improper disposal practices, contributes to air, water, and soil contamination driven by complex biological, chemical, and physical processes within landfills (Choudhary et al. 2024).

Engineers face a multitude of obstacles when dealing with contaminated soil, especially in the proximity to landfill sites. Contaminated soil can adversely affect the geotechni-

*Addresses of the authors:
Department of Civil Engineering, Delhi Technological University, New Delhi, India.
[1]Shivangi Bharadwaj (corresponsing author): shivangi.bharadwaj23@gmail.com, ORCID: https://orcid.org/0009-0006-1636-8856
[2]Ashok Kumar Gupta: akgupta@dtu.ac.in, ORCID: https://orcid.org/0009-0005-1344-3876
[3]Anil Kumar Sahu: sahuanilkr@yahoo.co.in, ORCID: https://orcid.org/0000-0003-3504-8090

© 2026 E. Schweizerbart'sche Verlagsbuchhandlung, Stuttgart, Germany
DOI: 10.1127/zdgg/0529
www.schweizerbart.de
1860-1804/0529 $ 7.65



Effect of Leachate Contamination on Soil

Shivangi Bharadwaj(✉), Ashok Kumar Gupta, and Anil Kumar Sahu

Department of Civil Engineering, Delhi Technological University, Delhi, India
shivangi.bharadwaj23@gmail.com

Abstract. Open dumping landfills are the most popular technique used to deal with Municipal solid waste in India. Landfills are a commonly used but also heavily polluted method of waste management. The improper design and maintenance of landfills have caused numerous environmental problems, such as the production of leachate and the emission of methane, which contributes to the increase in greenhouse gas emissions. Leachate is a liquid that enters a landfill from external sources such as rain. Delhi, India's capital, currently have three high-operating landfill sites which are Ghazipur, Okhla, and Bhalswa. The types of contaminants that are present in leachate depend on the nature of the solid waste as well as the concurrent physical, chemical, and biological processes taking place inside the landfill. Hazardous metals like Pb, Zn, and Cd participate in biogeochemical cycles, and rather than the elements' overall concentration, their mobility is mostly influenced by their chemical forms. Because of this, heavy metals in leachate may travel outside the confines of the disposal site and constitute a serious threat to the soil and water table nearby the landfill. In the present study, the Bhalswa dumpsite is selected, there is no liner system provided in this particular landfill. This study focuses on the change in characteristics and behavior of soil due to contamination from landfill leachate. The number of soil samples had been collected from the Bhalswa landfill to Delhi Technological University in lateral direction. The number of samples of leachate from Bhalswa landfill site has also been collected and estimated for chemical analysis. This chemical analysis includes parameters like pH, total dissolved solids, total hardness, alkalinity, sulphate, nitrate, electrical conductivity, chloride, and phosphate, calcium, magnesium.

Keywords: Landfill · Municipal solid waste · Leachate · physicochemical analysis · Soil contamination

1 Introduction

Municipal solid waste (MSW) quantity and quality are influenced by a number of variables, including population, way of life, dietary habits, standard of living, level of industrial and commercial activity, local cultural traditions, and climate. The seven most significant metro areas' MSW generation rates are expressed in terms of both daily and per-capita basis (CPCB 2004). The environment will be severely impacted if this much of large volume of MSW is not managed properly. Conflicts between environmental and developmental goals continue to plague waste management policies and practices.

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2025
A. Kashyap et al. (Eds.): ICCRIP 2023, LNCE 601, pp. 98–108, 2025.
https://doi.org/10.1007/978-981-96-4902-0_9

Characterization of Contaminated Soil and Groundwater Due to Landfill Leachate



Shivangi Bharadwaj, Ashok Kumar Gupta, and Anil Kumar Sahu

Abstract Landfills are the most prevalent waste disposal sites globally due to their cost-effectiveness. This study focuses on the Bhalswa dumpsite in Delhi, operational since 1992, which lacks scientific lining material, allowing continuous wastewater injection. The research involves a physicochemical analysis of soil, groundwater, and leachate samples from around the Bhalswa Landfill near Rohini, Delhi. Parameters examined include TDS, chloride content, TH, alkalinity, pH, calcium, and magnesium levels, along with sulfate, nitrate, and phosphate content in centrifuged soil samples. The study encompasses 10 soil samples, 10 groundwater samples, and 5 leachate samples collected from the Bhalswa dumpsite to the DTU campus, Rohini. Some soil samples exhibited a pH range of 8.0–8.4, indicating alkalinity. These results are compared with uncontaminated soil samples from the university area to assess longitudinal variations in physiochemical parameters and to benchmark against standard soil chemical content ranges. The findings show changes in soil behavior, and groundwater results are compared with the BIS drinking water standards, revealing significant pollution and unsafe conditions for drinking and household use near the Bhalswa landfill due to high levels of chemical contaminants.

Keywords Municipal solid waste · Landfill leachate · Physiochemical analysis · Soil contamination

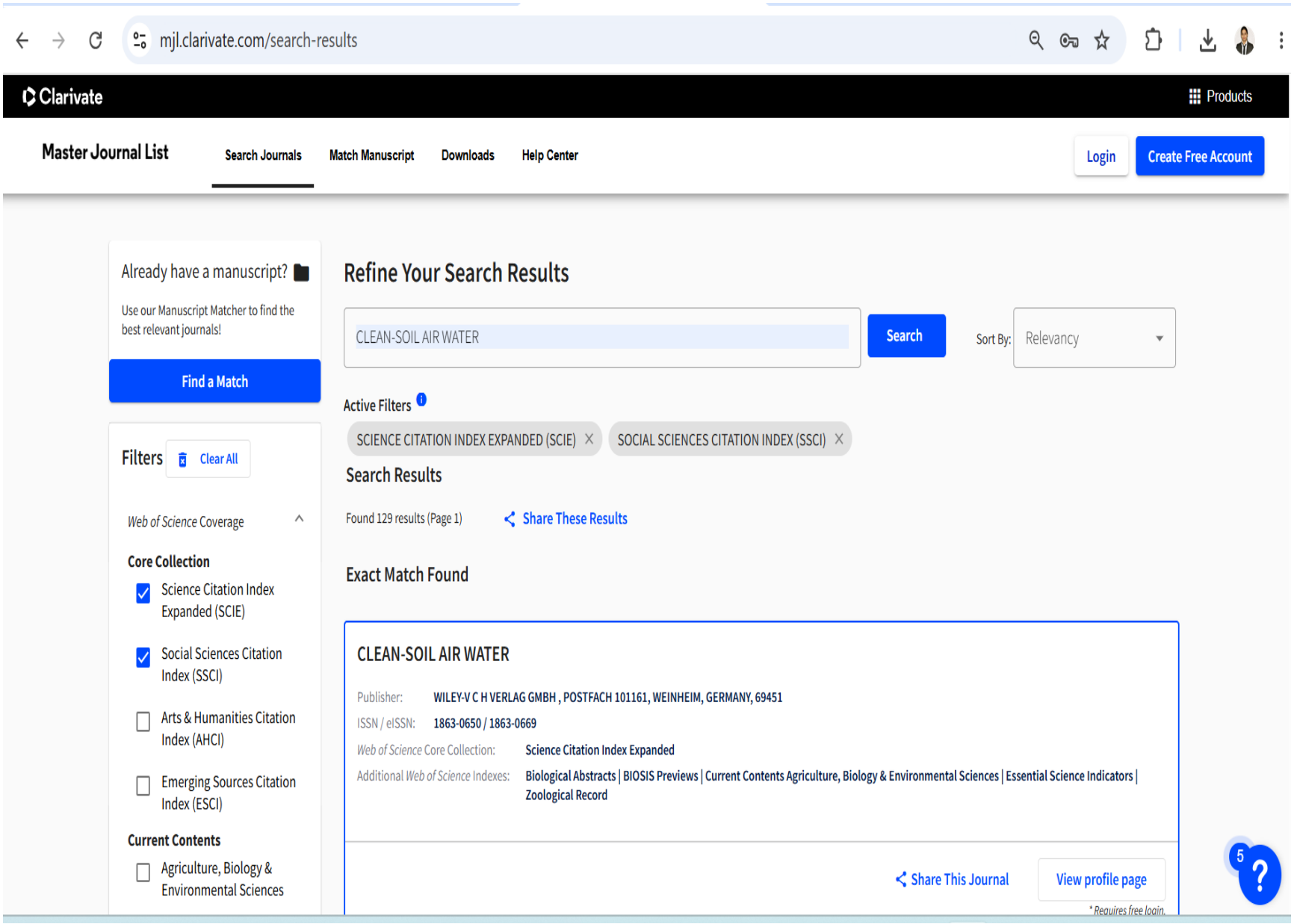
S. Bharadwaj (✉) · A. K. Gupta · A. K. Sahu
Department of Civil Engineering, Delhi Technological University, New Delhi, India
e-mail: shivangibharadwaj_2k21phdce03@dtu.ac.in

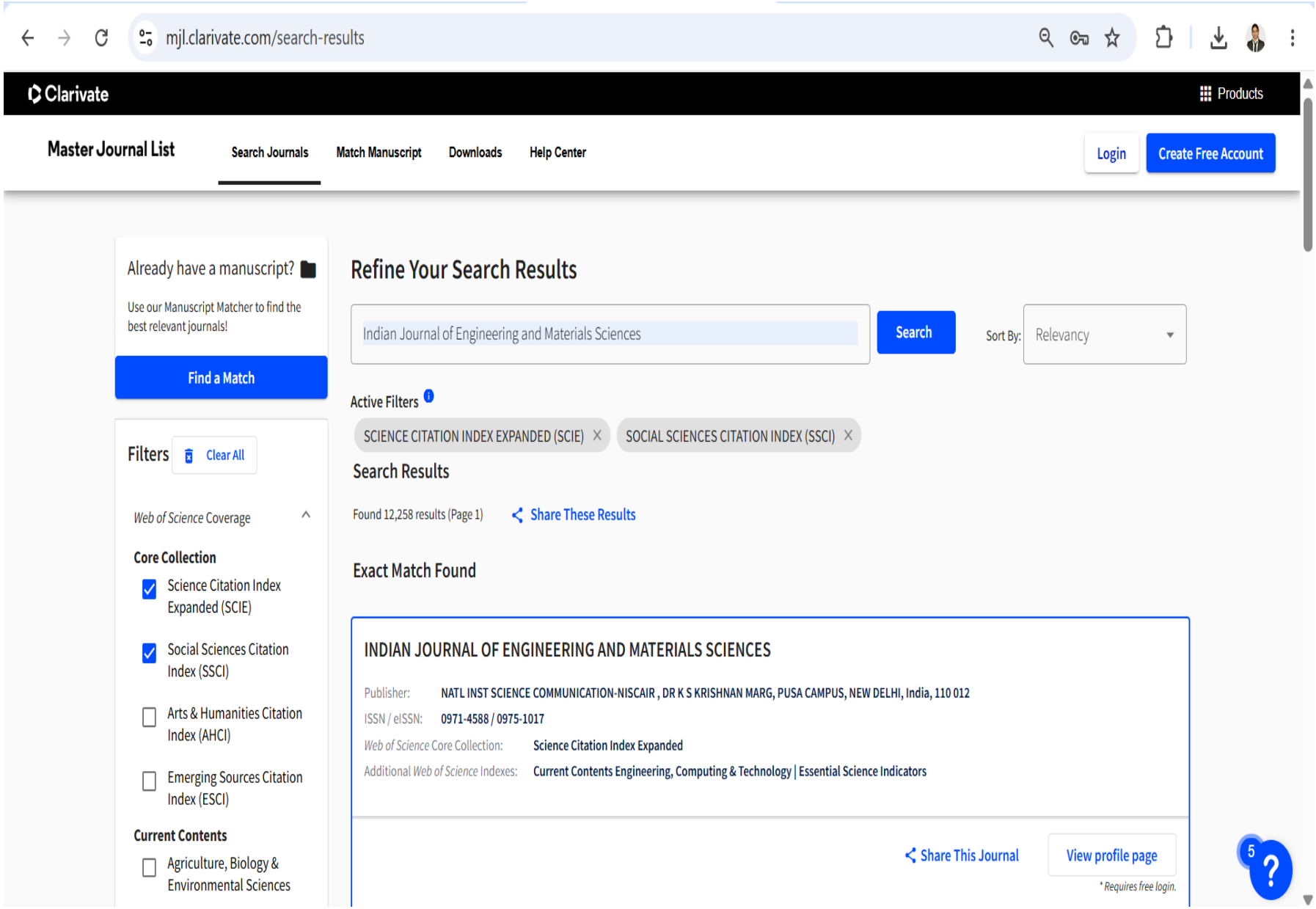
A. K. Gupta
e-mail: akgupta@dtu.ac.in

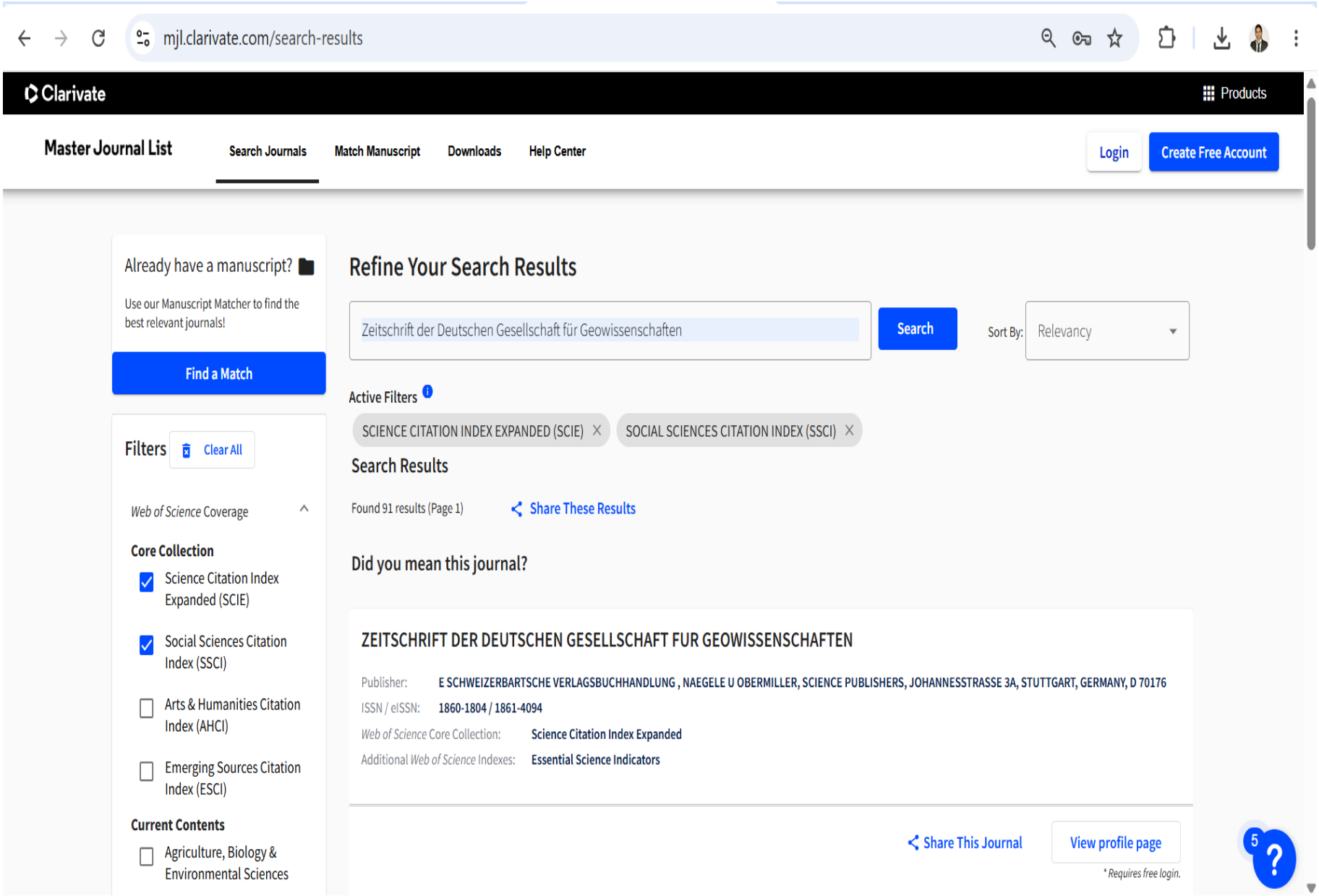
A. K. Sahu
e-mail: anilsahu@dtu.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2025

M. S. Dixit et al. (eds.), *Proceedings of the Indian Geotechnical Conference (IGC 2024), Volume 4*, Lecture Notes in Civil Engineering 700,
https://doi.org/10.1007/978-981-96-7771-9_7


mjl.clarivate.com/search-results
Clarivate
Products
Master Journal List
Search Journals
Match Manuscript
Downloads
Help Center
Login
Create Free Account
Already have a manuscript?
Use our Manuscript Matcher to find the best relevant journals!
Find a Match
Filters
Clear All
Web of Science Coverage
Core Collection
Science Citation Index Expanded (SCIE)
Social Sciences Citation Index (SSCI)
Arts & Humanities Citation Index (AHCI)
Emerging Sources Citation Index (ESCI)
Current Contents
Agriculture, Biology & Environmental Sciences
Refine Your Search Results
CLEAN-SOIL AIR WATER
Search
Sort By:
Relevancy
Active Filters
SCIENCE CITATION INDEX EXPANDED (SCIE)
SOCIAL SCIENCES CITATION INDEX (SSCI)
Search Results
Found 129 results (Page 1)
Share These Results
Exact Match Found
CLEAN-SOIL AIR WATER
Publisher: WILEY-V C H VERLAG GMBH , POSTFACH 101161, WEINHEIM, GERMANY, 69451
ISSN / eISSN: 1863-0650 / 1863-0669
Web of Science Core Collection: Science Citation Index Expanded
Additional Web of Science Indexes: Biological Abstracts | BIOSIS Previews | Current Contents Agriculture, Biology & Environmental Sciences | Essential Science Indicators | Zoological Record
Share This Journal
View profile page
* Requires free login.
5


mjl.clarivate.com/search-results
Clarivate
Products
Master Journal List
Search Journals
Match Manuscript
Downloads
Help Center
Login
Create Free Account
Already have a manuscript?
Use our Manuscript Matcher to find the best relevant journals!
Find a Match
Refine Your Search Results
Indian Journal of Engineering and Materials Sciences
Search
Sort By: Relevancy
Active Filters
SCIENCE CITATION INDEX EXPANDED (SCIE)
SOCIAL SCIENCES CITATION INDEX (SSCI)
Filters
Clear All
Search Results
Found 12,258 results (Page 1)
Share These Results
Web of Science Coverage
Core Collection
Science Citation Index Expanded (SCIE)
Social Sciences Citation Index (SSCI)
Arts & Humanities Citation Index (AHCI)
Emerging Sources Citation Index (ESCI)
Current Contents
Agriculture, Biology & Environmental Sciences
Exact Match Found
INDIAN JOURNAL OF ENGINEERING AND MATERIALS SCIENCES
Publisher: NATL INST SCIENCE COMMUNICATION-NISCAIR , DR K S KRISHNAN MARG, PUSA CAMPUS, NEW DELHI, India, 110 012
ISSN / eISSN: 0971-4588 / 0975-1017
Web of Science Core Collection: Science Citation Index Expanded
Additional Web of Science Indexes: Current Contents Engineering, Computing & Technology | Essential Science Indicators
Share This Journal
View profile page
* Requires free login.
5


mjl.clarivate.com/search-results
Clarivate
Products
Master Journal List
Search Journals
Match Manuscript
Downloads
Help Center
Login
Create Free Account
Already have a manuscript?
Use our Manuscript Matcher to find the best relevant journals!
Find a Match
Refine Your Search Results
Zeitschrift der Deutschen Gesellschaft für Geowissenschaften
Search
Sort By: Relevancy
Active Filters
SCIENCE CITATION INDEX EXPANDED (SCIE)
SOCIAL SCIENCES CITATION INDEX (SSCI)
Filters
Clear All
Web of Science Coverage
Core Collection
Science Citation Index Expanded (SCIE)
Social Sciences Citation Index (SSCI)
Arts & Humanities Citation Index (AHCI)
Emerging Sources Citation Index (ESCI)
Current Contents
Agriculture, Biology & Environmental Sciences
Search Results
Found 91 results (Page 1)
Share These Results
Did you mean this journal?
ZEITSCHRIFT DER DEUTSCHEN GESELLSCHAFT FUR GEOWISSENSCHAFTEN
Publisher: E SCHWEIZERBARTSCHE VERLAGSBUCHHANDLUNG , NAEGELE U OBERMILLER, SCIENCE PUBLISHERS, JOHANNESSTRASSE 3A, STUTTGART, GERMANY, D 70176
ISSN / eISSN: 1860-1804 / 1861-4094
Web of Science Core Collection: Science Citation Index Expanded
Additional Web of Science Indexes: Essential Science Indicators
Share This Journal
View profile page
* Requires free login.

NICMAR
UNIVERSITY
PUNE

7TH INTERNATIONAL CONFERENCE ON
CONSTRUCTION, REAL ESTATE, INFRASTRUCTURE AND PROJECT MANAGEMENT (ICCRIP)
11th and 12th August 2023

Certificate of Participation

PRESENTER
ICCRIP
2023

This is to certify that

Mr./Ms./Dr. Shivangi Bharadwaj

has presented a paper/case study titled

Effect of Leachate Contamination on Soil

in the 7th ICCRIP.

DR. RAJNI KANT RAJHANS	DR. RAJESH GOYAL	DR. SUSHMA S. KULKARNI	DR. ANIL KASHYAP
(Co-Convener)	(Convener)	(Conference Patron)	(Conference Chief Patron)

Knowledge Partners: IPM, InCoBAN, NIDM

Industrial Partner: STEINER

Platinum Sponsors: L&T Construction, ICICI Bank, BEXEL MANAGER, Verstraete, TATA PROJECTS

MIT
A group of Academic & Research Institutions
Chhatrapati Sambhajinagar

INDIAN GEOTECHNICAL CONFERENCE (IGC 2024)

Geotechnical Engineering For A Sustainable Tomorrow [GEST 2024]

MIT, Chhatrapati Sambhajinagar, (Aurangabad) - 431 010, Maharashtra, India

AURANGABAD CHAPTER

Certificate of Participation

This is to certify that, Shivangi Bharadwaj

from Delhi Technological University, Delhi *has presented a technical paper titled*

Characterisation of contaminated Soil & Groundwater due to landfill leachate *at*

the ***Indian Geotechnical Conference 2024****, organized at the MIT, Chhatrapati Sambhajinagar, Maharashtra, by the Indian Geotechnical Society, Aurangabad Chapter in association with the Department of Civil Engineering, MIT, between 19 - 21 December 2024.*

Dr. Manish S. Dixit
Organising Secretary,
IGC 2024

Dr. Uttam B. Kalwane
Chairman,
IGC 2024

Dr. Nilesh G. Patil
Director,
MIT

PEC
EXPLORE INNOVATE EXCEL

ASCE INDIA SECTION NORTHERN REGION

CERTIFICATE

This is to certify that

Ms. / Mr. / Dr. **Shivangi Bharadwaj**

has presented "**Landfill Mining And Waste Management: Challenges, Opportunities, And Innovations**"

at

ASCE International Conference

on

Challenges and Innovations for Sustainable Smart Cities (CISSC 2025)

(February 7–9, 2025, Hotel Novotel, Chandigarh, India)

organized by

ASCE India Section Northern Region (ASCE IS NR) & Punjab Engineering College, Chandigarh

Dr Har Amrit Singh Sandhu
Organising Chair

Prof. Sanjay Kr. Singh
Co-Patron

RECYCLE 2023

4th International Conference on Waste Management

CERTIFICATE OF PARTICIPATION

This is to certify that

Shivangi Bharadawaj

presented Oral Presentation on the topic **Analysis Of Landfill Slope Stability in Bhalswa, Delhi** in Recycle 2023, The 4th International Conference on Waste Management, Organised by Waste Management Research Group, Indian Institute of Technology Guwahati on 18th and 19th May 2023.

Meena Kh

Dr. Meena Khwairakpam

Convener

Prof. Ajay Kalamdhad

Convener

विद्याधनं सर्वधन प्रधानम्

INDIAN INSTITUTE OF TECHNOLOGY JAMMU

INTERNATIONAL CONFERENCE ON SUSTAINABLE MATERIALS AND INFRASTRUCTURE

CERTIFICATE

OF PARTICIPATION

This certificate is awarded to

Shivangi Bharadwaj

for presenting a paper titled "Natural Attenuation Enhancement of Uncontaminated Soil through Compost Amendment: A Column Study" in the ***International Conference on Sustainable Materials and Construction (ICSMI)*** *held from* ***16th to 18th December 2025*** *at* ***IIT Jammu****. We acknowledge his/her enthusiasm to learn about sustainable materials and construction practices.*

Dr. Deepak Yadav
Organising Secretary, ICSMI 2025

Prof. Anurag Misra
Chairperson, ICSMI 2025

Dr. Ankit Kathuria
Organising Secretary, ICSMI 2025

PLAGIARISM REPORT

Final thesis

ORIGINALITY REPORT

9%	6%	7%	2%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.jeeng.net Internet Source	<1%
2	Viraj Gunarathne, Ankur J. Phillips, Alessandra Zanoletti, Anushka Upamali Rajapaksha et al. "Environmental pitfalls and associated human health risks and ecological impacts from landfill leachate contaminants: Current evidence, recommended interventions and future directions", Science of The Total Environment, 2023 Publication	<1%
3	essay.utwente.nl Internet Source	<1%
4	everant.org Internet Source	<1%
5	www.preprints.org Internet Source	<1%
6	Bappa Mukherjee, Sohan Kar, Kalachand Sain. "Machine Learning Assisted State-of-the-Art-of Petrographic Classification From Geophysical Logs", Pure and Applied Geophysics, 2024 Publication	<1%
7	Submitted to Notre Dame de Namur University Student Paper	<1%
8	Al Sawafi, Ali Said Ali. "Developing an ESP Lifespan Predictive Model Using Artificial Intelligence: A Case Study on an Omani Oil Field", Sultan Qaboos University (Oman) Publication	<1%
9	Yahya Kooch, Azam Nouraei, Liping Wang, Xiang Wang, Donghui Wu, Rosa Francaviglia, Jan Frouz, Mohammad Kazem Parsapour. "Long-term landfill leachate pollution suppresses soil health indicators in natural ecosystems of a semi-arid environment", Chemosphere, 2024 Publication	<1%
10	researchonline.jcu.edu.au Internet Source	<1%

BRIEF CURRICULUM VITAE

Ms. Shivangi Bharadwaj is currently working as an Assistant Professor at the School of Planning and Architecture (SPA), New Delhi, and is pursuing her Doctor of Philosophy (Ph.D.) in Civil Engineering from Delhi Technological University (DTU), Delhi. Her doctoral research primarily focuses on the impact of landfill leachate on soil behaviour and groundwater quality, with particular emphasis on soil contamination, heavy metal migration, environmental geotechnics, and the application of machine learning techniques for predicting contamination levels.

She completed her B.Tech. in Civil Engineering from Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal, Madhya Pradesh, in 2014. Subsequently, she obtained her M.Tech. in Civil Engineering, specializing in Structural Engineering, from Madhav Institute of Technology and Science (MITS), Gwalior, in 2016.

Ms. Bharadwaj began her academic career as an Assistant Professor at Bharat Institute of Engineering and Technology, Hyderabad, from September 2016 to November 2017. She subsequently served as an Assistant Professor at Kommuri Pratap Reddy Institute of Technology, Hyderabad, from December 2017 to December 2018. She later worked as a University Fellow at The NorthCap University, Gurugram, from July 2019 to July 2021, before joining Delhi Technological University for her doctoral research in 2021.

Her teaching and research interests include Geotechnical Engineering, Environmental Engineering, Soil and Groundwater Contamination, Structural Engineering, and the application of Machine Learning in Civil Engineering.