

# **Ecological assessment of vegetation dynamics of the Aik stream with reference to the Industrial Pollution in Sialkot, Pakistan**



**Ujala Ejaz**

(PhD)

**Department of Plant Sciences,**

**Faculty of Biological Science**

**Quaid-i-Azam University**

**Islamabad, Pakistan**

**2025**



# **Ecological assessment of vegetation dynamics of the Aik stream with reference to the Industrial Pollution in Sialkot, Pakistan**



*A thesis submitted to the Quaid-i-Azam University in partial fulfillment of the requirements for the  
Degree of Doctor of Philosophy*

**In**

**Botany/Plant Sciences**

**(Plant Ecology and Conservation)**

**By**

**Ujala Ejaz**

**Department of Plant Sciences,**

**Faculty of Biological Science**

**Quaid-i-Azam University**

**Islamabad, Pakistan**

**2025**




QUAID-I-AZAM UNIVERSITY ISLAMABAD  
DEPARTMENT OF PLANT SCIENCES

24 June 2024

**Subject: Author Deceleration**

I Ms. Ujala Ejaz hereby declared that my PhD thesis entitled “**Ecological assessment of vegetation dynamics of the Aik stream with reference to the Industrial Pollution in Sialkot, Pakistan**” is my own work and no part of this thesis has been previously submitted to this or any other University as part of the requirement for a higher degree. The contents of this thesis are the results of my own work unless otherwise acknowledged in the text or by reference.

At any time, if my statement is found to be incorrect even after my graduation, the university has the right to withdraw my PhD degree.

  
\_\_\_\_\_  
Ms. Ujala Ejaz



**QUAID-I-AZAM UNIVERSITY ISLAMABAD**  
**DEPARTMENT OF PLANT SCIENCES**

24 June 2024

**Subject: Plagiarism Undertaking**

It is hereby declared that the research work presented in this dissertation titled “**Ecological assessment of vegetation dynamics of the Aik stream with reference to the Industrial Pollution in Sialkot, Pakistan**” is my own work with no significant contribution from any other person. Small contributions or help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero-tolerance policy of the HEC and Quaid-i-Azam University, Islamabad towards plagiarism. Therefore, I as an author of the above titled thesis declared that no portion of this thesis has been plagiarized, and any material used as reference is properly referred/cited.

I undertake that if I am found faulty of any formal plagiarism in the above titled thesis even after award of PhD degree, the university reserves the right to withdraw/ revoke my PhD degree. HEC and the University have the right to publish my name on the HEC/ University website on which the name of the student is placed who submitted plagiarized thesis.



---

**Student / Author Signature**





## QUAID-I-AZAM UNIVERSITY

Faculty of Biological Sciences,

Department of Plant Sciences,

Islamabad, PAKISTAN

### SIMILARITY INDEX CERTIFICATE

It is certified that **Ms. Ujala Ejaz** has completed his PhD research work and compilation of the thesis. The title of his thesis "**Ecological assessment of vegetation dynamics of the Aik stream with reference to the Industrial Pollution in Sialkot, Pakistan**" has been checked on Turnitin for similarity index and found 12% that lies in the limit provided by HEC (19%).

**Dr. Shujaul Mulk Khan**

Associate Professor

Department of Plant Sciences,

Quaid-i-Azam University Islamabad



### Certificate of Approval

This is to certify that the research work presented in this thesis, entitled "**Ecological assessment of vegetation dynamics of the Aik stream with reference to the Industrial Pollution in Sialkot, Pakistan**" was conducted by **Ms. Ujala Ejaz** under the supervision of Professor **Dr. Shujaul Mulk Khan**. No part of this thesis has been submitted anywhere else for any other degree. This thesis is submitted as a partial fulfillment of the requirement for the degree of Doctor of Philosophy in the field of **Plant Sciences/Botany** (Plant Ecology and Conservation), to the Department of Plant Sciences, Quaid-i-Azam University Islamabad Pakistan.

Student Name : **Ms. Ujala Ejaz**

Signature: \_\_\_\_\_



#### Examination Committee:

External Examiner 1:

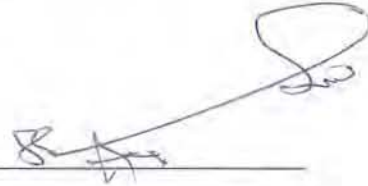
**Prof. Dr. Sheikh Saeed Ahmad**

Chairperson

Environmental sciences

Fatima Jinnah Women University

Signature: \_\_\_\_\_



External Examiner 2:

**Dr. Sundus Javed**

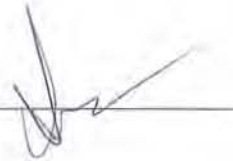
Associate Professor

Department of Biosciences,

COMSATS University,

Park Road, Chak Shahzad, Islamabad, Pakistan

Signature: \_\_\_\_\_



Internal Examiner/Supervisor:

**Prof. Dr. Shujaul Mulk Khan**

Professor

Department of Plant Sciences,

Quaid-i-Azam University, Islamabad, Pakistan

Signature: \_\_\_\_\_



Chairman:

**Prof. Dr. Hassan Javed Chaudhary**

Professor and Chairperson

Department of Plant Sciences,

Quaid-i-Azam University, Islamabad, Pakistan

Signature: \_\_\_\_\_



Dated:

08/01/2025



## AL QURAN

Assuredly the creation of the heavens and the earth is a greater (matter) than the creation of men: yet most men understand not. Not equal are the blind and those who (clearly) see: not are (equal) those who believe and work deeds of righteousness, and those who do evil. Little do ye learn by admonition! The Hour will certainly come therein is no doubt: yet most men believe not and your Lord says: Call on me; I will answer your (Prayer): but those who are too arrogant to serve me will surely find themselves in Hell in humiliation! It is Allah who has made the Night for you, that ye may rest therein, and the Day, as that which helps (you) to see. Verily Allah is full of Grace and Bounty to men: yet most men give not thanks. Such is Allah, your Lord, the Creator of all things. There is no God but He: then how ye are deluded away from the truth! Thus, are deluded those who want to reject the Signs of Allah.

*Al Mumin Verses 40:57 to 63.*



## Dedication

*This thesis is dedicated to my beloved family,  
whose unwavering love, sacrifices, and encouragement have  
been the foundation of my strength and perseverance.*

*To my mentors and teachers,  
whose guidance, wisdom, and belief in my potential have  
inspired me to aim higher and think deeper.*

*To my friends,  
for their patience, understanding, and constant support  
throughout this journey.*

*And to the countless researchers and visionaries,  
whose tireless pursuit of knowledge continues to illuminate  
the path for generations to come.*



## ACKNOWLEDGEMENTS

I am profoundly grateful to the All-Mighty Allah, the Source of Knowledge and Wisdom, for granting me the intellect and fortitude to embark on this academic journey. My heartfelt appreciation extends to the esteemed Holy Prophet Muhammad (PBUH), whose teachings have provided invaluable guidance and illumination throughout this endeavor. Completing this thesis stands as a testament to the abundant blessings bestowed upon me by Allah throughout my life journey. I am deeply indebted to all those who supported me in various capacities during the completion of this manuscript. Their contributions, both large and small, have played a pivotal role in shaping this research.

Foremost, I extend my sincerest gratitude to my supervisor, **Dr. Shujaul Mulk Khan**, Associate Professor of Plant Sciences at Quaid-i-Azam University Islamabad, whose unwavering guidance and encouragement propelled me towards the successful completion of my PhD project. Additionally, I express my gratitude to Prof. Dr. Mushtaq Ahmad, Chairman of the Department of Plant Sciences, for his invaluable support.

I am indebted to **Prof. Jens-Christian Svenning** for his meticulous review of my thesis, offering invaluable insights and endless inspiration. His dedication to research and science has profoundly motivated me, and I aspire to emulate his clarity and simplicity in writing throughout my career.

Special thanks are extended to **Dr. Noreen Khalid**, Assistant Professor at the University of Sialkot, for her consistent assistance and guidance throughout my master's and PhD journey.

I am deeply grateful to **HEC, Pakistan**, for their invaluable support of my studies in Denmark through the International Research Support Initiative Program (IRSIP). Additionally, I sincerely appreciate the generous financial assistance provided by HEC, Pakistan, through the "Access to Scientific Instrumentation Program" at GC University Lahore, Pakistan.

I am immensely grateful to my friends and colleagues, **Sadia Jahangir** and **Qurat ul Ain**, for their unwavering support and encouragement during my doctoral journey. Additionally, I extend my heartfelt thanks to my PhD seniors, **Abdullah** and **Zeeshan Ahmed**, for their invaluable guidance and support. Furthermore, I appreciate **Shah Fahad** for his expertise in data analysis.



To my lab fellows at the Plant Ecology and Conservation Lab, I offer my deepest gratitude for their camaraderie, empathy, and unwavering moral support, which have been indispensable throughout this journey.

Lastly, I express profound admiration and gratitude to my parents, **Ejaz Ahmed Virk** and **Rahat Ashraf**, for their unwavering love, prayers, and unwavering support. I also pay tribute to my late **Nani Balkees Akhter**. Their sacrifices and encouragement have served as the cornerstone of my success. I am also thankful to my siblings, **Huma**, **Mani**, **Mehak**, and **Kashaf**, for their steadfast support and grounding presence.

I offer my heartfelt regards and blessings to all my loved ones who supported me along the way. I apologize for not being able to mention each one individually. Your support has been invaluable, and I am deeply grateful.

**Ujala Ejaz**



## **Publications arising to date from this Dissertation.**

The following papers have been published based on some results presented in the thesis:

1. Ejaz U, Khan SM, Aqeel M, Khalid N, Sarfraz W, Naeem N, Han H, Yu J, Yue G and Raposo A (2022) Use of *Parthenium hysterophorus* with synthetic chelator for enhanced uptake of cadmium and lead from contaminated soils—a step toward better public health. *Front. Public Health* 10:1009479. doi: 10.3389/fpubh.2022.1009479 **(Article Published)**
2. Ejaz, U., Khan, S. M., Khalid, N., Ahmad, Z., Jehangir, S., Fatima Rizvi, Z., & Raposo, A. (2023). Detoxifying heavy metals: a multipronged study of tolerance strategies against heavy metals toxicity in plants. *Frontiers in Plant Science*, 14, 1154571 **(Article Published)**.
3. Ujala Ejaz, Shujaul Mulk Khan, Sadia Jehangir, Zeeshan Ahmad, Abdullah Abdullah, Majid Iqbal, Noreen Khalid, Aisha Nazir and Jens-Christian Svenning (2024). Monitoring the Industrial waste polluted stream - Integrated analytics and machine learning for water quality index assessment, *Journal of Cleaner Production*, Volume 450, 2024, 141877, ISSN 0959-6526 **(Article Published)**.
4. Ejaz, U., Khan, S. M., Khalid, N., Jehangir, Ain, A., Z. & Abdullah, A. (2024). DYNAMIC INTERPLAY BETWEEN INDUSTRIAL POLLUTION AND FLORISTIC COMPOSITION: AN ECOLOGICAL PERSPECTIVE. *Pakistan Journal of Botany* **(Article Published)**.
5. Ejaz, U., Khan, S. M., Shah, S. F. A., Khalid, N., Jehangir, S., Rizvi, Z. F., & Svenning, J.-C. (2025). Advanced Integrative Analytics for Data-Driven Risk Assessment of Ecological and Human Health Risks from Heavy Metal Contaminations in Soil. *Journal of Hazardous Materials Advances*, 100596 **(Article Published)**.
6. Ejaz, U., Khan, S. M., Khalid, N., Jehangir, S., Shah, S. F. A., & Svenning, J. C. (2024). Elucidating the phytoremediation potentials and ecophysiological mechanisms of indicator plants in the industrial polluted region. *Journal of Environmental Management*, 366, 121821 **(Article Published)**.
7. Ejaz, U. (2024). Ujala Ejaz, Shujaul Mulk Khan, Muhammad Aqeel, Noreen Khalid, Wajiha Sarfraz, Nayab Naeem, Heesup Han\*, Jongsik Yu, Gong Yue and António Raposo. *The One Health Approach in the Context of Public Health*, 47 **(Book Chapter Published)**.



8. Ujala Ejaz, Shujaul Mulk Khan, Noreen Khalid, and Abdullah (2021). Ecological assessment of the vegetation along the upper Aik Stream (Nallah Aik), Sialkot Pakistan. Abstract accepted in 4th international conference on Biosciences (ICBS-2021) Organized by Biological Society of Pakistan (BSP) 15th June 2021, at Pakistan Academy of sciences **(Abstract Published)**.
9. Ujala Ejaz, Noreen Khalid, Zeeshan Ahmad, Abdullah Abdullah, Shujaul Mulk Khan (2022). Distribution of *Parthenium hysteroporous* in relation with the environmental variables of the Sialkot Industrial Zone, Pakistan. Abstract accepted for oral presentation (online) in 4th Euro-Mediterranean Conference for Environment Integration, 20-23 October 2022, SOUSSE TUNISIA, Under ID: EMCEI-22-P17 and for publication in the conference proceedings by Springer **(Abstract Published)**.
10. Ujala Ejaz, Shujaul Mulk Khan, Noreen Khalid, Abdullah and Sadia Jehangir (2022) Estimation of Plant biodiversity and indicator species for biomonitoring of leather industrial polluted zone of Sialkot, Pakistan, Abstract accepted for oral presentation at Botanical Society of Pakistan, 8th international and 17th National Conference on “Advances in Plant Science in the Era of Climate Change” at Government College University, Lahore & University of the Punjab, Lahore, Pakistan **(Abstract Published)**.
11. Ejaz, U., Khan, S. M., Khalid, N., Jehangir, & Svenning, J.C. (2024). Alien Species Importance in Riparian Vegetation. Ecological Indicators **(Under Review)**.



## Table of Contents

<b>1 Chapter</b> .....	3
<b>General Introduction</b> .....	3
1.1 Introduction .....	3
1.2 Study area .....	9
1.3 Hypothesis .....	10
1.4 Aim of the study .....	11
1.5 Objectives .....	11
1.6 Thesis Structure .....	11
<i>Review Paper Published From 1st Chapter</i> .....	13
<b>2 Chapter</b> .....	14
<b>Water Pollution Assessment in Aik-Stream: Leveraging Advanced Analytical and Machine Learning Techniques</b> .....	14
2.1 Introduction .....	14
2.1.1 Water pollution by industries .....	14
2.1.2 The Health Impacts of Water Pollution .....	15
2.1.3 Rising Concerns of Industrial Water Pollution in Pakistan .....	15
2.1.4 Industrial Pollution in Sialkot, Pakistan .....	16
2.1.5 The Aik-Stream of Sialkot, Pakistan .....	17
2.1.6 Machine learning as a novel monitoring tool .....	18
2.1.7 Problem and Novelty Statement .....	19
2.1.8 Hypothesis: .....	20
2.1.9 Objectives: .....	20
2.2 Materials and Methods .....	21
2.2.1 Study Area .....	21
2.2.2 Study Design .....	23
2.2.3 Data collection .....	24
2.2.4 Sampling sites (S1-S51) .....	24
2.2.5 Sampling sites (S51-S117) .....	26
2.2.6 Sampling sites (S117-150) .....	29
2.2.7 Analytical Procedures in the Laboratory .....	30
2.2.8 Statistical analysis .....	34
2.2.8.1 Hierarchical Agglomerative Cluster Analysis (HACA) .....	35
2.2.8.2 Analysis of Variance (ANOVA) .....	35



2.2.8.3 Principal Component Analysis (PCA).....	35
2.2.9 Innovative Monitoring Approaches Using Machine Learning Models.....	36
2.2.10 The Water Quality Index (WQI) .....	36
2.2.11 Machine learning models.....	38
2.2.11.1 Machine Learning (ML) Research Methodology.....	38
2.2.11.2 Evaluation Metrics.....	39
2.2.11.3 Feature selection .....	40
2.3 Results .....	41
2.3.1 Classification of the sampling sites .....	44
2.3.2 Variation among the zones and within the parameters.....	48
2.3.3 Principal Component Analysis (PCA).....	55
2.3.4 Innovative Monitoring with Machine Learning Models .....	60
2.3.4.1 Calculation of WQI .....	60
2.3.5 Evaluating ML models .....	61
2.3.6 Identifying Optimal Input Combinations .....	64
2.4 Discussion .....	66
2.5 Conclusion.....	73
<i>Paper Published from 2nd Chapter .....</i>	<i>74</i>
<b>3 Chapter .....</b>	<b>75</b>
<b>Assessing Soil Pollution around Aik-Stream: Utilizing Advanced Methods for Ecological Risk Assessment .....</b>	<b>75</b>
3.1 Introduction .....	75
3.1.1 Soil Pollution by Industrial Wastewater: A Global Issue.....	76
3.1.2 Heavy Metals in Soil: An Escalating Environmental Concern .....	77
3.1.2.1 Chromium (Cr) .....	77
3.1.2.2 Cadmium (Cd).....	78
3.1.2.3 Copper (Cu).....	78
3.1.2.4 Arsenic (As) .....	79
3.1.2.5 Lead (Pb) .....	80
3.1.2.6 Mercury (Hg).....	80
3.1.2.7 Nickel (Ni).....	81
3.1.2.8 Zinc (Zn).....	81
3.2 Materials and Methods .....	83
3.2.1 Study Design.....	83
3.2.2 Analytical Procedures in the Laboratory .....	84



3.2.3 Statistical analysis.....	85
3.2.3.1 The selected indices.....	85
3.2.3.2 Potential Ecological Risk Index (PERI) .....	86
2.2.4 Heavy Meatal Absorption from Water to Soil .....	87
3.2.5 Self-Organizing Map (SOM).....	88
3.2.5.1 SOM neural network learning step .....	88
3.2.5.2 Determination of neuron size in the SOM competitive layer .....	90
3.2.5.3 Determination of the best classification number .....	90
3.2.6 Methodology of Structural Equation Models (SEM) .....	91
3.3 Results .....	92
3.3.1 Evaluation of Soil Parameters .....	92
3.3.2 Heavy Metal Contamination in the Soil .....	93
3.3.3 Self-Organizing Map (SOM).....	96
3.3.4 Cluster Analysis Using Davies-Bouldin Index (DBI) .....	99
3.3.5 Ecological risks assessment of heavy metals in soil.....	101
3.3.6 Identifying Key Factors Influencing Heavy Metals' Ecological Risk via Structural Equation Modeling .....	102
3.4 Discussion .....	114
3.5 Conclusion.....	118
<i>Research Article Published from 3<sup>rd</sup> Chapter.....</i>	120
<b>4 Chapter.....</b>	121
<b>Understanding Vegetation Dynamics: Analyzing Structure, Composition, Distribution Patterns, Identifying Indicator Plant Species and the Role of Alien Species in Riparian Zone.....</b>	121
4.1 Introduction .....	121
4.1.1 Impact of Anthropogenic Activities on Natural Vegetation.....	122
4.1.2 Indicator Species and Their Significance .....	122
4.1.3 Anthropogenic Disturbances Drive Alien and Invasive Plant Dominance .....	124
4.2 Material and Methods.....	126
4.2.1 Vegetation sampling .....	126
4.2.2 Physiological attributes.....	127
4.2.2.1 Density.....	127
4.2.2.2 Relative density .....	127
4.2.2.3 Frequency .....	128
4.2.2.4 Relative frequency .....	128
4.2.2.5 Cover .....	128



4.2.2.6 Relative cover .....	128
4.2.2.7 Basal area.....	128
4.2.2.8 Cover classes for herb and shrub species .....	129
4.2.3 Importance Value Index (IVI) .....	129
4.2.4 Climatic Data.....	129
4.2.5 Soil Data Collection and Analysis.....	129
4.2.6 Statistical Analysis .....	130
4.2.6.1 Species-area curves.....	131
4.2.6.2 Cluster Analysis.....	131
4.2.6.3 Indicator Species Analysis (ISA) .....	131
4.2.6.4 Diversity Indices, Species richness and evenness .....	132
4.2.6.4.1 Shannon Index (H').....	132
4.2.6.4.2 Simpson Index (D).....	132
4.2.6.4.3 Simpson Index of Diversity (1-D).....	133
4.2.6.4.4 Pielou's Evenness (J) / Species Evenness .....	133
4.2.6.4.5 Species richness.....	133
4.2.6.4.6 Brillouin Diversity Index .....	133
4.2.6.4.7 Margalef Diversity Index.....	133
4.2.6.4.8 Equitability-J Index .....	134
4.2.6.4.9 Fisher's Alpha ( $\alpha$ ).....	134
4.2.6.4.10 Berger-Parker Index .....	134
4.2.6.4.11 Chao-1 Estimator.....	134
4.2.6.4.12 iChao-1 Estimator.....	135
4.3 Results .....	135
4.3.1 Plant species composition.....	135
4.3.2 Species area curves .....	142
4.3.3 Abundant and rare plants of the Aik-Stream .....	143
4.3.3.1 Abundant and rare trees .....	143
4.3.3.2 Abundant and rare shrubs .....	144
4.3.3.3 Abundant and rare herbs.....	145
4.3.4 Vegetation Classification.....	147
4.3.5 Species Richness and Diversity Indices .....	148
4.3.5 Indicator of Vegetation zones of the Aik-Stream region.....	149
4.3.5.1 Less Polluted Zone (LPZ).....	149
4.3.5.2 High Polluted Zone (HPZ) .....	150



4.3.5.3 Moderately Polluted Zone .....	151
4.3.6 Direct Ecological Gradient through Canonical Correspondence Analysis (CCA).....	154
4.3.7 Impact of Environmental Factors on Plant Abundance Using GLM .....	155
4.3.7.1 Less Polluted Zone (LPZ).....	155
4.3.7.2 High Polluted Zone (HPZ) .....	156
4.3.7.3 Moderate Polluted Zone (MPZ) .....	158
4.3.8 Correlation between Alien and Invasive with Native Species Diversity.....	159
4.3.9 Factors Shaping Species Composition .....	160
4.4 Discussion .....	161
4.5 Conclusion.....	165
<i>Research Articles Published from 4<sup>th</sup> Chapter .....</i>	<i>167</i>
<b>5 Chapter.....</b>	<b>169</b>
<b>Harnessing Nature's Remedies: Assessing Phytoremediation Potential and Physiological Responses of Indicator Plant Species .....</b>	<b>169</b>
5.1 Introduction .....	169
5.1.2 Phytoremediation: A Green Approach to Environmental Cleanup .....	170
5.1.3 Mechanisms of Heavy Metal Uptake by Plants.....	171
5.1.4 Factors Affecting Heavy Metal Accumulation in Plants.....	172
5.1.5 Advantages of Phytoremediation.....	173
5.1.6 Challenges in Phytoremediation.....	174
5.1.7 Physiological Responses of Plants to Heavy Metal Pollution.....	175
5.2 Material and Methodology .....	177
5.2.1 Quantification of heavy metals.....	177
5.2.2 Phytoremediation examination .....	178
5.2.3 Determination of photosynthetic pigments.....	178
5.2.4 Proline analysis.....	179
5.2.5 Statistical analysis.....	179
5.2.5.1 Ordinary Least Square (OLS).....	179
5.2.5.2 Probit Model.....	180
5.2.5.3 Generalized Mixed Effect Model (GLMM) .....	180
5.3 Results .....	181
5.3.1 Vegetation along the polluted stream .....	181
5.3.2 Soil of the Polluted Ecosystem .....	182
5.3.3 Phytoremediation of heavy metals .....	183
5.3.4 Plant physiological responses .....	190



5.3.5 Chlorophyll-a.....	191
5.3.6 Chlorophyll-b.....	192
5.3.7 Carotenoids.....	193
5.3.8 Proline.....	194
5.3.9 Phytoremediation Efficacy in Indicator Plants: Insights from Probit and OLS Regression Models .....	199
5.3.10 Generalized Mixed-Effect Modeling Approach .....	205
5.4 Discussion .....	214
5.5 Conclusion.....	218
<i>Research Article Published from 5<sup>th</sup> Chapter .....</i>	<i>219</i>
<b>6 Chapter.....</b>	<b>220</b>
<b>Monitoring Land Use Dynamics: Remote Sensing and NDVI Analysis for Assessing Urban-Industrial Expansion and Ecological Impacts in Aik-Stream .....</b>	<b>220</b>
6.1 Introduction .....	220
6.1.1 Land Use and Land Cover Change (LULC).....	220
6.1.2 Urban-industrial Development and Riparian Areas .....	221
6.1.3 Mapping urban-industrial development.....	221
6.1.4 Analysis Techniques.....	222
6.1.5 Statement of the Problem and Scope of the Study .....	223
6.2. Material and Methodology .....	224
6.2.1 Research Design .....	224
6.2.2 Data Acquisition .....	225
6.2.3 Image preprocessing .....	226
6.2.4 Image classification and training sample collection.....	226
6.2.5 Normalized Difference Vegetation Index (NDVI).....	227
6.2.6 Accuracy Assessment .....	227
6.2.7 Kappa Statistics .....	228
6.3 Results .....	229
6.3.1 LULC Change: Every Consecutive Year, 1998-2023 .....	230
6.3.2 Accuracy Assessment .....	234
6.3.3 NDVI .....	236
6.4 Discussion .....	238
6.5 Conclusions .....	241
<b>7 Chapter .....</b>	<b>243</b>
<b>Discussion and Synthesis .....</b>	<b>243</b>



7.1. Conclusion.....	249
7.2 Recommendations and Strategies for Enhancing Stream and Catchment Area.....	250
7.3 For the improvement of stream water quality .....	251
7.4 For the protection of natural floral diversity .....	252
7.5 Future Prospects .....	252
References .....	254



## List of Figures

Figure 1.1 Geographical positioning of sialkot .....	10
Figure 2. 1 propose model working daigram.....	20
Figure 2. 2 map depicting the industrialized city of sialkot in pakistan.....	23
Figure 2. 3 (A-D) illustrates the upstream area of the aik-stream.....	26
Figure 2. 4 (A-D) illustrates the mid-stream area of the aik-stream.....	29
Figure 2. 5 (A-B) illustrates the downstream area of the aik-stream.....	30
Figure 2. 6 (A-J) some glimpse of the laboratory work.....	34
Figure 2. 7 dendrogram showing different clusters of sampling zones .....	45
Figure 2. 8 dendrogram showing the concentration of parameters in different sites.....	46
Figure 2. 9 heavy metal in three zoneS (LPZ, MPZ, AND HPZ).....	48
Figure 2. 10 physicochemical characteristics of Aik stream water quality in three zones .....	48
Figure 2. 11 (A-S) Box plot variance comparison in LPZ, MPZ, AND HPZ).....	55
Figure 2. 12 PCA biplot depicting the three distant zones of the Aik stream.....	59
Figure 2. 13 scree plot for PCA component selection .....	60
Figure 2. 14the top-performing input variables (4th and 13th) for WQI.....	62
Figure 2. 15 RF algorithm's best inputs (2nd and 10th) for WQI.....	64
Figure 3.1 Map illustrating the industrial city of Sialkot.....	84
Figure 3.2 Basic structure of the som network .....	89
Figure 3.3 A conceptual model illustrating.....	91
Figure 3.4 Insights into the degree and distribution of heavy metal contamination.....	94
Figure 3. 5 This visualization shows the content of soil elements using SOM. ....	98
Figure 3.6 Minimum, mean and maximum values of soil heavy metal.....	101
Figure 3.7 Assessment of ecological risk from various heavy metals.....	102
Figure 3.8 (A-H) Illustrations depict a linear structural equation model.....	106
Figure 4.1 Percentage distribution of herbs, shrubs and trees around aik-stream .....	136
Figure 4.2 The species-area curves of 182 plant species .....	143
Figure 4.3 The topmost abundant tree species around the aik-stream, sialkot, pakistan.....	144
Figure 4.4 The rare tree species around the aik-stream, sialkot, pakistan. ....	144
Figure 4.5 The dominant shrubs species around the aik-stream, sialkot, pakistan.....	145
Figure 4.6 The rare shrubs species around the aik-stream, sialkot, pakistan.....	145
Figure 4.7 The top ten most abundant herb species.....	146
Figure 4.8 Rare herb species with minimum IVI.....	146
Figure 4.9 Cluster analysis (CA) of the aik-stream .....	147
Figure 4.10 TWCA dendrogram depicted plant species.....	148
Figure 4.11 Indicator species of all the studied zones .....	155
Figure 5.1 chlorophyll-a of the identified indicators plant species .....	192
Figure 5. 2 Chlorophyll-b content (mg/g) of the identified indicators plant .....	193
Figure 5. 3 Carotenoids of indicators plant species present in LPZ, MPZ and HPZ.....	194
Figure 5.4 Proline indicators plant species present in LPZ, MPZand HPZ.....	195
Figure 5.5 Physiological impact on indicator plants from heavy metal accumulation.....	205
Figure 6.1 NDVI and land use and land cover data collection,.....	225
Figure 6.2 Illustrates the collective alterations in land use and land cover .....	231
Figure 6.3 Presents the comprehensive transformations in LULC .....	232
Figure 6. 4 Illustrates the overall changes in land use and land cover.....	232
Figure 6.5 Displays the comprehensive alterations in land use and land cover .....	233
Figure 6. 6 Illustrates the comprehensive changes in land use and cover. ....	233
Figure 6.7 A brief study of changes in land use and cover.....	234
Figure 6.8 Normalized difference vegetation index (NDVI) of the study area .....	237



Figure 6. 9 Normalized difference vegetation index (NDVI).....	238
--	-----



## List of Tables

Table 2. 1 Description and measurement techniques of the explanatory variables.....	32
Table 2. 2 Classification of ccme-wqi and corresponding water status.....	38
Table 2. 3 Models developed from the different combinations of the water parameters .....	39
Table 2.4 Descriptive statistics of water quality parameters of the aik-stream .....	41
Table 2.5 Pearson correlation matrix of all studied variables.....	43
Table 2.6 Zone-wise descriptive statistics among all parameters of the aik stream. ....	47
Table 2.7 Mean deviation of the water quality in different sites of the aik stream. ....	50
Table 2.8 Normality tests for all the studied water quality parameters. ....	51
Table 2.9 Total variance explained of the water quality parameters in the aik stream.....	57
Table 2.10 The contribution of the six components of pca in aik stream water quality. ....	58
Table 2.11 Evaluation measures for the gb algorithm in WQI.....	62
Table 2.12 Evaluation measures for the rf algorithm in WQI. ....	63
Table 2.13 Displays the results of a regression analysis aimed at identifying WQI. ....	65
Table 2.14 Multicollinearity statistics analysis for wqi parameters .....	66
Table 3.1 The different type of model of the soil contamination.....	62
Table 3.2 Summary statistics of soil samples from the study area .....	93
Table 3.3 The trasfer rate of the heavey metal from water to soil .....	96
Table 3.4 Contamination factor and Degree of contamination factor.....	111
Table 3.5 Potential ecological risk (PER) and potential ecological risk index (PERI) .....	110
Table 4.1 braun blanquet cover classes with their mid points for herb and shrub species....	129
Table 4.2 Detail of plant species along with their habit. ....	136
Table 4.3 Diversity indices of all the three vegetation zones of the aik-stream. ....	149
Table 4.4 The topmost indicator species.....	150
Table 4. 5 The topmost indicator species.....	150
Table 4.6 The foremost three indicators species.....	151
Table 4.7 Other indicator species of the three identified zones. ....	151
Table 4.8 Generalize linear model to determine the impact of environmental on plants .....	156
Table 4.9 Indicate the model fit value for GLM. ....	156
Table 4.10 Generalize linear model to determine the impact of environmental on plants ....	157
Table 4.11 Indicate the model fit value for GLM. ....	157
Table 4.12 Generalize linear model to determine the impact of environmental on plants ....	158
Table 4. 13 Indicate the model fit value for GLM. ....	159
Table 4.14 Spearman's correlation between native diversity and alien/invasive species. ....	165
Table 4.15 Mental test.....	180
Table 5.1 The identified indicator plant species.....	181
Table 5.2 Summary statistics of soil samples from the study area .....	182
Table 5.3 BCF, TF and BAC for different heavy metals of indicator species.....	184
Table 5.4 The phytoremediation capacity of identified indicator plant species .....	187
Table 5.5 Detailed description of chlorophyll-a, b and total carotenoids.....	193
Table 5.6 Summary statistics. ....	196
Table 5.7 Ordinary least square (ols) and probit model.....	198
Table 5.8 Generalized mixed effect model .....	208
Table 6.1 Classification of the satellite images for study area during 1998 to 2023.....	226
Table 6.2 Land use and cover changes in sialkot, pakistan (1998-2023). ....	230



Table 6.3 Changes in LULC classes from 1998 to 2023 in the study area.....	231
Table 6.4 The accuracy of landsat-4, 5 and 8 OLI imagery .....	236



## List of Appendix

Appendix table 1. Plant species list with their families and phytogeographic elements.....	286
Appendix table 2 calculated ccme-wqi along with water status of the aik-stream.....	293
Appendix table 3 Contamination factor (CF) and degree of contamination (DC) .....	296
Appendix table 4 (PER) and (PERI).....	298
Appendix table 5 Summary statistics (LSEM) for the Cu .....	300
Appendix table 6 Summary statistics of the SEM fit values for Cu. ....	300
Appendix table 7 Summary statistics of (LSEM) for Cr. ....	301
Appendix table 8 Summary statistics of the SEM fit values for Cr.....	301
Appendix table 9 Summary statistics of the (LSEM) for Zn, .....	302
Appendix table 10 Summary statistics of the SEM fit values for Zn. ....	302
Appendix table 11 Summary statistics of the (LSEM) for Cd.....	303
Appendix table 12 Summary statistics of the SEM fit values for Cd. ....	303
Appendix table 13 Summary statistics of the (LSEM) for Ni. ....	304
Appendix table 14 Summary statistics of the SEM fit values for Ni.....	304
Appendix table 15 Summary statistics of the (LSEM) for Pb .....	305
Appendix table 16 Summary statistics of the SEM fit values for Pb.....	305
Appendix table 17 Summary statistics of the (LSEM) for As .....	306
Appendix table 18 Summary statistics of the SEM fit values for As. ....	306
Appendix table 19 Summary statistics of the (LSEM) For Hg.....	307
Appendix table 20 Summary statistics of the SEM fit values for Hg.....	307
Appendix table 21 Average NDVI and NDWI of the study area during (1998-2023).....	308
Appendix figure 1 GB models for training and testing.....	308
Appendix figure 2 The graphical depiction of the random forest (RF).....	311



## Abbreviations

AI	Artificial Intelligence
ML	Machine Learning
GB	Gradient Boost
RF	Random Forest
AdB	Adaptive Boosting
SVR	Support Vector Regression
BR	Bayesian Regression
K-NN	K-Nearest Neighbors
TWCA	Two-way Cluster Analysis
ISA	Indicator Species Analysis
SAC	Species Area Curves
SOM	Self-Organizing Map
PERI	Potential Ecological Risk Index
CCME	Canadian Council of Ministers of the Environment
WQI	Water Quality Index
IS	Imperial Surface
DCA	Detrended Correspondence Analysis
CCA	Canonical Correspondence Analysis
NDVI	Normalized Difference Vegetation Index
SEM	Structural Equation Modeling
GDP	Gross Domestic Product
LNMDS	Local Nonmetric Multidimensional Scaling
NCBI	National Center for Biotechnology Information
GIS	Geographical Information System
IVI	Importance Value Index
RF	Relative Frequency
RD	Relative Density
RC	Relative Cover
EC	Electrical Conductivity
TDS	Total Dissolved Solids
OM	Organic matter
CaCO <sub>3</sub>	Calcium Carbonate
K	Potassium
P	Phosphorus
Mn	Manganese
Ni	Nickle
Cd	Cadmium
Cr	Chromium
Cu	Copper
Zn	Zinc
Fe	Iron
Mg	Magnesium
Ca	Calcium
As	Arsenic
Hg	Mercury
CTAB	Cetyltrimethylammonium Bromide
ITS	Internal Transcribed Spacer
BLAST	Basic Local Alignment Search Tool



OLS	Ordinary Least Square
AIC	Akaike Information Criteria
GFI	Goodness of Fit Statistic
AGFI	Adjusted Goodness of Fit Statistic
CFI	Comparative Fit Index
SRMR	Standardized Root Mean Square Residual
RMSEA	Root Mean Square Error of Approximation
NFI	Normed Fit Index
NNFI	Non-normed Fit Index
PTEs	Potential Toxic Elements
ROS	Reactive Oxygen Species
HPZ	Highly Polluted Zone
LPZ	Less Polluted Zone
MPZ	Moderately Polluted Zone
BAC	Bioaccumulation Coefficient
TF	Translocation Factor
BCF	Biological Concentration Factor



## Abstract

The Aik-Stream area in Sialkot, Pakistan, confronts many environmental challenges from industrial activities and urban expansion. These challenges encompass water quality degradation, soil pollution, loss of plant biodiversity, and land use changes, necessitating urgent attention and comprehensive management strategies. This thesis thoroughly investigates these issues, employing a multidisciplinary approach that integrates spatial analysis, machine learning, statistical modeling, ecological techniques, and remote sensing/GIS technologies. The overarching objective is to provide insights, solutions, and recommendations for sustainable environmental management in the Aik-Stream area. This research aims to assess the environmental health of the Aik-Stream area and develop strategies for mitigating pollution and fostering ecosystem restoration. Specific objectives include assessing water and soil quality to pinpoint pollution hotspots and understand contamination extents, investigating plant biodiversity and elucidating its correlation with environmental factors, evaluating the phytoremediation potential of pollution-tolerant plant species, and analyzing land use/land cover changes to gauge urban expansion impacts on ecological integrity. The methodology encompasses several key steps, starting with spatial analysis to categorize the stream into pollution severity zones and outlining areas of heightened contamination. Machine learning models, specifically Gradient Boosting (GB) and Random Forests (RF), predict the Water Quality Index (WQI) more effectively. Soil pollution is evaluated using a blend of techniques, including Self-Organizing Map (SOM), Potential Ecological Risk Index (PERI), and Structural Equation Modeling (SEM), to discern influential factors. Plant biodiversity is analyzed via statistical techniques such as General Linear Model (GLM), Canonical Correspondence Analysis (CCA), and Indicator Species Analysis (ISA) to identify key species and their ecological preferences. Quantitative Ecological Techniques (QET) and physiological response analyses assess the phytoremediation potential of pollution-tolerant plant species. Remote sensing and GIS technologies track land use/land cover changes over time, uncovering urban expansion patterns and their ecological repercussions. The analysis uncovers significant exceedances of water quality parameters and heavy metals in soil, especially in industrial-influenced zones. Machine learning models exhibit precision in WQI prediction, promising enhanced water quality monitoring. Soil pollution assessments underscore heavy metal threats, predominantly from industrial and agricultural activities. Plant biodiversity analysis identifies vital species and their intricate relationship with environmental variables. Effective plant species for phytoremediation are identified, offering avenues for pollution control and ecosystem restoration. Land use/land cover change analysis reveals substantial urban expansion, necessitating informed urban planning to mitigate ecological degradation. The findings underscore the urgent need for effective pollution control policies, soil remediation strategies, and informed urban planning to combat environmental degradation and safeguard ecological integrity. Integrated approaches, blending scientific research, stakeholder engagement, and policy formulation, are imperative for achieving environmental sustainability in the Aik-Stream area and beyond. In conclusion, this research provides valuable insights and methodologies to tackle environmental challenges in the Aik-Stream area. Sustainable management practices, informed by scientific research and stakeholder collaboration, are pivotal for pollution mitigation and ecosystem resilience. Concerted efforts are essential to implement recommended strategies, monitor their efficacy, and adapt management practices to evolving environmental dynamics. Future research should prioritize implementing recommended strategies, monitoring their effectiveness, and adapting management practices to evolving environmental conditions. Collaboration between stakeholders, policymakers, and researchers remains pivotal for attaining sustainable environmental management goals.



and ensuring long-term ecological health in the Aik-Stream area and similar regions worldwide.



## **General Introduction**

### **1.1 Introduction**

Water drives significant global physical, chemical, and biological changes (Yang, Yang, & Xia, 2021). The world's oceans and rivers account for approximately 97% of Earth's water reserves, with freshwater resources comprising the remaining 3% (Khilchevskiy & Karamushka, 2021). Regarding the distribution of freshwater, 68.7% was stored in glaciers and ice caps at the poles. 30.1% exists as groundwater, 0.3% is found in surface waters, and the remaining 0.9% is present in different forms (Inglezakis, Pouloupoulos, Arkhangelsky, Zorpas, & Menegaki, 2016). Water is a crucial component of the biosphere and an essential element for living organisms (Ball, 2008). The absence of water for more than a few days is life-threatening, and even a limited supply can significantly alter the distribution patterns of wildlife and humans (Organization, 2005). Freshwater, a finite and indispensable resource, is vital for the sustenance of living organisms and the execution of human endeavors, including agriculture, industry, and domestic activities (R. K. Mishra, 2023). The utilization of freshwater resources is intertwined with the genesis of human civilizations (B. K. Mishra, Kumar, Saraswat, Chakraborty, & Gautam, 2021). Water's significance extends beyond mere survival, contributing profoundly to the evolution of human societies. Civilizational development historically emerged in regions rich in quality and quantity of freshwater. Proximity to freshwater sources, particularly rivers, was a common characteristic of ancient civilizations (Angelakis et al., 2021). Renowned early societies such as the Mesopotamian, Egyptian, Chinese, and Indo-Gangetic civilizations flourished along riverbanks, utilizing these waterways for domestic needs, agriculture, and irrigation. Rivers like the Euphrates, Tigris, Nile, Yangtze, Indus, and Ganges were not just geographical features but vital lifelines that nurtured and sustained these ancient cultures (Best, 2019).

Freshwater resources are broadly categorized into three types: lotic (encompassing rivers and streams), lentic (including lakes and ponds), and groundwater (found in aquifers). Rivers and streams are distinguished by their unidirectional flow, typically exhibiting velocities greater than 0.1 m/sec (D. C. Martin, 1988; J. L. Martin, Schottman, & McCutcheon, 2018). In their pristine state, these lotic systems support stable aquatic ecosystems, but they are increasingly degraded due to overexploitation driven by human



demands (Rockström et al., 2014). The concern over freshwater as a resource has recently intensified, mainly due to the surge in global population, which has placed unprecedented strain on water resources (Arnell, 1999). This situation has been exacerbated by the industrial revolution and the rapid expansion of urban areas, further compounding the stress on these vital resources (Bogardi et al., 2012).

Rapid industrialization, extensive urbanization, intensive agricultural practices, and burning fossil fuels are key anthropogenic activities that have significantly altered the natural state of aquatic ecosystems (Muruganandam et al., 2023). These human activities have led to the degradation of stream and river ecosystems (Peters & Meybeck, 2000), impacting water quality and affecting the structure and function of aquatic biota (Meybeck, 2004). Streams and rivers worldwide face numerous environmental challenges due to anthropogenic influences within their catchment areas (Oyedotun & Ally, 2021). The pollution of freshwater resources is a global concern, particularly acute in developing and underprivileged countries. While developed nations have implemented various technologies to protect waterways from pollution, the intensity of aquatic pollution remains critical in developing countries (Matta, 2010), where effluents are often discharged into rivers and streams without adequate environmental safeguards (Kaushal et al., 2021). Most developing countries lack the infrastructure to control water pollution and enforce water quality standards (Radelyuk, Tussupova, Klemeš, & Persson, 2021). Both natural and anthropogenic factors influence the water quality of aquatic ecosystems, which is crucial for supporting biological communities and meeting human needs (Muruganandam et al., 2023). Surface water quality information is vital for assessing pollution and the long-term environmental impacts of human activities in a region (Yan, Shen, & Zhou, 2022). Pristine aquatic systems typically show less variation in water quality parameters than polluted ones (Tay, 2021). Human activities in catchment areas can introduce a variety of pollutants, including heavy metals, organic pollutants, nutrients, salts, and synthetic compounds (Akhtar, Syakir Ishak, Bhawani, & Umar, 2021), which degrade water quality and disrupt the ecological integrity of lotic ecosystems (Pearson, Connolly, Davis, & Brodie, 2021). Streams and rivers passing through urban and industrial areas experience severe degradation, with significant losses in biotic integrity, rendering many of these freshwater resources unsafe for human consumption (Bănăduc et al., 2022). Among lotic systems, streams are more vulnerable to anthropogenic activities compared to rivers, as they often undergo extreme fluctuations in discharge (Allan, Castillo, & Capps, 2021). Urban streams are particularly susceptible, showing irregular discharge, altered water



chemistry, and disrupted food chains (Mohanavelu, Shrivastava, & Naganna, 2022). The spatial and temporal variations in water quality of streams are influenced by both natural and human activities in the catchment area (Hamid, Bhat, & Jehangir, 2020). The impact of pollutants is typically greatest near their source, but tends to dilute over distance (Ryan, 2020). Streams affected by industrial and urban activities exhibit more pronounced spatial variations in water quality (Hamid et al., 2020).

Numerous studies have highlighted the adverse effects of both natural and anthropogenic factors on water quality, analyzed on a spatial and temporal basis worldwide (Akhtar et al., 2021; Alimonti, Mariani, Prodi, & Ricci, 2022; Han et al., 2020; Talbot & Chang, 2022; Ustaoglu, Tepe, & Tas, 2020). Seasonal natural processes influencing water quality include rainfall, weathering of parent rock materials, erosion, surface runoff, sediment transport, and changes in stream hydrology and flow (Akhtar et al., 2021). In contrast, anthropogenic factors encompass a range of activities impacting water quality, varying in intensity and duration, originating from both point and non-point sources (Fang, Deitch, Gebremicael, Angelini, & Ortals, 2024). Point sources are where pollutants are directly discharged into surface water, such as untreated industrial effluents and domestic sewage (A. Tariq & Mushtaq, 2023). These are more prevalent in urban and industrialized areas and are particularly problematic in regions where environmental regulations are not strictly enforced, often in developing countries (Gyurkovich & Gyurkovich, 2021). Non-point sources disperse pollutants across various locations, such as through surface runoff from urban and agricultural areas following rainfall (Wang, Fu, Qiao, Bi, & Liu, 2023). Pollutants like atmospheric deposits from fossil fuel combustion and industrial emissions can be washed into streams during rainfalls (Zeng et al., 2022). Similarly, agriculture fertilizers, pesticides, and soil improvers can leach into groundwater or enter streams and rivers with surface runoff (Craswell, 2021). The pollutant concentrations in surface runoff vary, typically higher in industrial areas than uninhabited regions (Zhu, Cheng, Li, Niu, & Wen, 2022). Identifying non-point sources is challenging due to their diffuse nature.

Heavy metals are a significant group of contaminants affecting surface water quality, originating naturally from parent rock material through weathering or anthropogenic activities (Akhtar et al., 2021). While alkali metals (Na, K) and alkaline earth metals (Ca, Mg) are abundant in the Earth's crust, heavy metals occur in trace amounts. Some heavy metals like Fe, Cr, Ni, Cu, and Zn are essential for various physiological processes in living organisms (Yu et al., 2021) but become toxic when concentrations exceed safe thresholds



(Budnicka, Sobiech, Kolmas, & Luliński, 2022). Non-essential metals are toxic even in trace amounts and have no known role in metabolic functions. The toxic effects of heavy metals vary within the food chain, becoming more pronounced at higher trophic levels among aquatic organisms (Nkwunonwo, Odika, & Onyia, 2020).

Water pollution affects aquatic environments and significantly impacts adjacent terrestrial ecosystems through leaching and precipitation (Akhtar et al., 2021). This interaction is particularly pronounced during monsoon seasons when increased water flow facilitates the spread of contaminated water across land areas (Chowdhury, 2010). As a result, pollutants such as heavy metals and industrial residues infiltrate the soil, altering its composition and potentially disrupting its natural biogeochemical cycles. The consequences of this pollutant seepage are multifaceted (Gautam, Mishra, & Agrawal, 2021). The affected soil may experience compromised fertility and overall health. The accumulation of pollutants poses threats to plant life, potentially inhibiting growth and physiological processes (Alengebawy, Abdelkhalek, Qureshi, & Wang, 2021).

Furthermore, the absorption of these contaminants by plant roots raises concerns regarding bioaccumulation and the subsequent entry of these toxins into terrestrial and aquatic food chains. This situation underscores the broader implications for environmental health and food safety, with the consumption of crops irrigated with contaminated water presenting tangible risks to human health due to the elevated levels of heavy metals in the soil, which can impact plant functions and human well-being (Pullagurala et al., 2018). Polluted water transfers pollutants to adjacent land areas, predominantly through leaching and precipitation processes, particularly accentuated during monsoon seasons (Gupta, 2020). Increased water flow during these periods facilitates the spread of contaminated water over land, leading to heavy metals and industrial waste infiltration into nearby soils. This infiltration alters soil composition and can disrupt its natural biogeochemical cycles (Rahel & Olden, 2008). The implications of such seepage are multifaceted, affecting soil fertility and health and posing threats to plant growth and physiological functions (Chaudhry & Sidhu, 2022). The absorption of pollutants by plants raises concerns about bioaccumulation and the entry of these toxins into terrestrial and aquatic food chains. This scenario underscores the risks to human health from consuming crops irrigated with contaminated water, thereby raising broader issues related to environmental health and food security. Consistent use of heavy metal-laden water for irrigation can increase the concentration of these metals in the



soil, adversely affecting plant health and posing significant human health risks (Hembrom, Singh, Gupta, & Nema, 2020).

Furthermore, the degradation of water quality results in the loss of biological communities near streams. Plants serve as effective biological indicators of water pollution in rivers and streams, responding sensitively to physical and chemical changes in the water, impacting biological communities. The ecological health of an ecosystem can be evaluated using plant indicators based on their presence or absence, relative abundance, and community structure and function (Sharma et al., 2011). Aquatic and terrestrial organisms, particularly plants, are crucial for understanding the impacts of human activities on stream ecosystems. Plants exhibit varied tolerance levels to pollution stress and respond to changes in an ecosystem's physical, chemical, and biological conditions caused by human actions (Vezzani et al., 2018). Local physical characteristics and regional interactions influence the distribution of plant species (Wisz et al., 2013). They are sensitive to human disturbances such as industrial effluent inflow, municipal waste, changes in stream discharge, habitat alteration, and fragmentation (Chakraborty, 2021). Sensitive plant species more accurately indicate stream health compared to tolerant species. However, tolerant species can adapt to unfavorable conditions created by natural or anthropogenic factors and recover quickly once the stress is reduced (Platts, Megahan, & Minshall, 1982). Identifying indicator plant species is increasingly crucial for assessing environmental health (Carignan & Villard, 2002). Most indicator species endure only a limited range of environmental pollutants, making them suitable for gauging the health of natural ecosystems. Conversely, rare species with narrow ecological tolerance are overly sensitive to pollution and are less reliable in reflecting pollution response.

On the other hand, ubiquitous species, with their broad tolerance, tend to be less responsive to pollution. Various environments feature numerous indicator species; for instance, lichens and bryophytes are frequently used to assess air pollution, as documented by (Aprile, Catalano, Migliozi, & Mingo, 2011; Conti & Cecchetti, 2001). Using indicator species offers a different approach from traditional chemical and physical methods of assessing environmental quality. Plant-based indicators offer a unique temporal perspective, mirroring organisms' lifespan or occupancy duration in specific habitats. Their utility is especially notable in detecting pollutants at low concentrations that typically require expensive and extensive analysis (Aschemann-Witzel, Gantriis, Fraga, & Perez-Cueto, 2021). The tolerance of these indicator species provides a simple and cost-effective means to gauge



pollution levels in various environments. These species, intricately linked with the complexities of ecosystem dynamics, provide an in-depth understanding of these constantly evolving natural settings. Monitoring these indicators in their natural habitats yields critical insights into physical and chemical shifts, ecological functions, and biodiversity variations (Noss, 1990).

Despite the challenging and toxic conditions pollutants create, certain plant species exhibit exceptional resilience, thriving in such environments. This resilience is not just a survival mechanism; it offers key insights into natural pollution mitigation strategies. Leveraging these adaptations could lead to innovative, biologically-based approaches to combat heavy metal pollution, aligning with sustainable environmental management goals (Ijaz et al., 2023). Plants capable of enduring high levels of heavy metals act as vital pollution indicators in specific habitats or biomes (Patra, Acharya, Pradhan, & Patra, 2021). Some of these plants possess an extraordinary ability to absorb or break down heavy metals at levels much higher than those in standard environments, serving as natural purifiers and playing a crucial role in the remediation of polluted sites (Kvesitadze, Sadunishvili, & Kvesitadze, 2009). Additionally, these plants have developed sophisticated phytoremediation strategies to withstand and mitigate the effects of their polluted environments. This adaptation underscores these species' significant role in environmental restoration and pollution control efforts (Sarwar et al., 2017).

The study of spatial distribution or mapping represents a crucial technique in geography and environmental statistics, focusing on organizing and analyzing phenomena across the Earth's surface (Bishop, James, Shroder Jr, & Walsh, 2012). This methodology is pivotal in summarizing and interpreting raw data directly or through more complex forms, offering a comprehensive view of diverse data aspects via graphical representations. The strategic use of colour differentiation often enhances this. A significant application of spatial data lies in its ability to identify areas with elevated pollution risks, thereby guiding policymakers toward regions that require urgent attention and intervention (Goodchild et al., 1996). Central to this approach is the geostatistical method, renowned for providing unbiased estimates of variable values in areas where direct sampling is not feasible (Haining & Haining, 1993). Complementing these methods, the study also explores the use of vegetation health measures, particularly the Normalized Difference Vegetation Index (NDVI). The NDVI, a robust tool, quantitatively assesses vegetation cover and biomass based on surface reflectance (Pettorelli et al., 2011). This index, which operates within the visible and infrared



spectrum (Huang, Tang, Hupy, Wang, & Shao, 2021), is acclaimed for its effectiveness in monitoring regional vegetation.

Moreover, its relevance extends to gauging vegetation productivity and assessing pollution impacts on vegetation (Becker & Choudhury, 1988). Integrating Geographic Information Systems (GIS) with remote sensing and advanced computational tools emerges as a formidable strategy for land-based pollution assessment. GIS provides a robust framework for storing, analyzing, and visualizing spatial data, which, when coupled with insights from remote sensing, culminates in a highly effective methodology for environmental analysis (Yin, Udelhoven, Fensholt, Pflugmacher, & Hostert, 2012).

## **1.2 Study area**

Sialkot is located in a region of Pakistan that falls between latitudes 32.240°N and 32.370°N as longitudes 73.590°E and 75.020°E (Figure 1.1) (Junaaid, Hashmi, & Malik, 2016). The city is located at a height of 244 meters above sea level. It has a population density of 903 people per square kilometer (Malik, Jadoon, & Husain, 2010). The climate in this area is known for its humid summers and winters and an average yearly rainfall of approximately 1,000 mm. Most rain falls during the monsoon season, forming deposits on the flat plains (A. Qadir & Malik, 2009). The alluvial soils here are of recent quaternary origin, predominantly composed of loamy and silty loam soils (Malik, Jadoon, et al., 2010). Over the past decade, Sialkot city has undergone rapid industrialization, urbanization, and agricultural development, making it highly susceptible to environmental pollution (Khalid et al., 2021b). The city is renowned globally for producing leather goods, sporting equipment, processed food items, ceramics, and surgical instruments. It houses approximately 92 tanneries, 244 leather garment and product manufacturing units, more than 900 sports manufacturing factories, 57 units dedicated to husking rice, and 14 mills producing flour (A. Qadir & Malik, 2009). Unfortunately, there is inadequate disposal of municipal and industrial waste in the region, leading to the unregulated discharge of solid waste and effluents from industries directly or indirectly into agricultural land, trenches, ponds, and natural watercourses (Naeem et al., 2021).

The primary focus of this study centers on the terrestrial area along the Aik-Stream, which falls within the geographical coordinates of latitude 32.630°N to 32.990°N and longitude 74.450°E to 74.690°E (A. Qadir & Malik, 2009). This stream holds significant importance as a major tributary of the Chenab River, traversing the city of Sialkot, Pakistan (Figure 1.1). It



spans 131.6 km, with a 315 Cs annual flow rate and a catchment area covering approximately 1,062 km<sup>2</sup>. Typically, the stream receives its lowest discharge during the early summer, while the utmost occurs during the monsoon season (Mahmood et al., 2014). The ecological balance in this region has been notably disrupted due to pollution from the industrial zone in Sialkot. As the stream flows through the city, it receives a substantial influx of wastewater, including toxic chemicals and heavy metals from municipal waste and industries, which drain into the Chenab River without treatment. This stream receives 52 million liters of wastewater per day, with an additional 1.1 million units of waste from leather-producing factories (Daily, 2006; Pakistan, 2007; A. Qadir & Malik, 2009; Abdul Qadir, Riffat Naseem Malik, & Syed Z. Husain, 2008b). The leather industry waste discharge includes organic and inorganic substances, toxic materials, i.e., heavy metals, chemically synthesized tannins, oils, resins, bio-toxins, and disinfectants (Garai, 2014; Maqbool et al., 2018; Rabelo et al., 2018; S. R. Tariq, Shaheen, Khalique, & Shah, 2010).

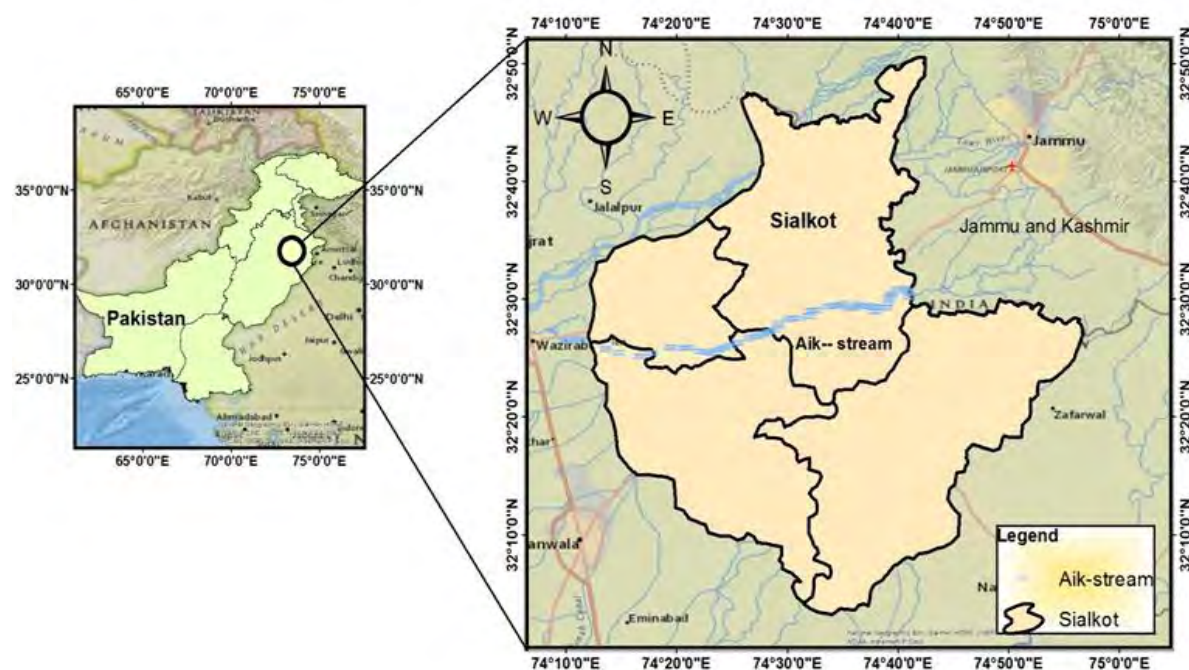


Figure 1.1 Geographical positioning of Sialkot and the Aik-stream in Pakistan.

### 1.3 Hypothesis

1. Elevated levels of pollutants from industrial and urban sources lead to decreased water quality in the Aik-Stream area compared to non-impacted regions.
2. The soil in the Aik-Stream area exhibits higher concentrations of heavy metals and other pollutants due to industrial activities and urbanization, resulting in degraded soil quality.



3. Pollution levels negatively impact plant diversity, with higher pollution associated with reduced species richness and abundance.
4. Pollution-tolerant plant species demonstrate higher phytoremediation abilities, effectively reducing pollutant levels in soil and water.
5. Remote sensing data analysis reveals significant land use changes over time, indicating the extent of urban expansion and its correlation with ecological degradation.

#### **1.4 Aim of the study**

This thesis endeavors to evaluate the environmental health of the Aik-Stream area in Sialkot, Pakistan, by examining water and soil quality, plant biodiversity, and land use changes. Through comprehensive analysis, the aim is to devise effective strategies for mitigating pollution, particularly from industrial activities, and promoting ecosystem restoration. The research seeks to provide actionable insights and recommendations for sustainable environmental management in the region by integrating multidisciplinary approaches, including spatial analysis, machine learning, and ecological techniques.

#### **1.5 Objectives**

1. Conduct a comprehensive assessment of water quality parameters in the Aik-Stream area to identify sources of pollution and determine the extent of contamination.
2. Evaluate soil quality and assess the presence of pollutants, particularly heavy metals, to understand their impact on ecosystem health.
3. Investigate the diversity and distribution of plant species in the Aik-Stream area to identify key indicator species and assess the ecosystem's overall health.
4. Assess the effectiveness of phytoremediation techniques using pollution-tolerant plant species to mitigate soil and water pollution in targeted areas.
5. Analyze land use and land cover changes over time using remote sensing and GIS technologies to understand the impact of urban expansion on ecological integrity.
6. Develop comprehensive management strategies based on the findings to address environmental challenges in the Aik-Stream area, including pollution control measures, restoration initiatives, and sustainable land use practices.

#### **1.6 Thesis Structure**

The thesis is structured into seven chapters, each addressing various environmental challenges in the Aik-Stream area of Sialkot, Pakistan. The first chapter serves as a general



introduction, providing an overview of the region and highlighting its many environmental issues. These challenges, stemming from industrial activities and urban expansion, encompass water quality degradation, soil pollution, loss of plant biodiversity, and land use changes. Recognizing the urgency of addressing these issues, the introduction outlines the overarching objective of the research - to provide insights, solutions, and recommendations for sustainable environmental management in the Aik-Stream area. Specific objectives include assessing water and soil quality, investigating plant biodiversity, evaluating phytoremediation potential, and analyzing land use changes. Subsequent chapters delve into specific aspects of the environmental challenges. Chapter two focuses on water pollution, exploring the causes and extent of water quality degradation in the Aik-Stream area. Through spatial analysis and machine learning models, it discusses pollution hotspots and contamination extents. Chapter three examines soil pollution, evaluating soil quality and contamination sources such as industrial and agricultural activities. Soil pollution assessment techniques, including Self-Organizing Map (SOM) and Potential Ecological Risk Index (PERI), are applied to discern influential factors. Chapter four focuses on vegetation dynamics, analyzing vegetation structure, composition, and regional distribution patterns. The correlation between vegetation dynamics and environmental factors is investigated to gain insights into ecosystem health. In chapter five, the phytoremediation potential of indicator plant species is assessed, along with their physiological responses in polluted environments. Effective plant species for pollution mitigation and ecosystem restoration are identified, offering potential solutions to environmental degradation. Chapter six delves into remote sensing and NDVI analysis to monitor land use/land cover changes. Urban expansion patterns and their ecological impacts are examined using remote sensing techniques. Finally, chapter seven provides a comprehensive discussion and synthesis of the findings from each chapter. Insights are integrated to propose recommendations for sustainable environmental management in the Aik-Stream area, addressing fundamental challenges and outlining future research directions.





# OPEN ACCESS

EDITED BY  
Muhammad Musq Khan,  
Zhejiang University, China

REVIEWED BY  
Mohammed Samir,  
Sulaiman Chaudhary University of Al-Qadisiyah,  
Iraq  
Mohammad Karaman Khan,  
Huazhong Agricultural University, China

\*CORRESPONDENCE  
Shujaul Mulk Khan  
✉ shujaulmulk@gmail.com  
Linda Heejung Lho  
✉ ltheethog@gmail.com  
Heesup Han  
✉ ltheesup.han@gmail.com

RECEIVED: 30 January 2023  
ACCEPTED: 06 April 2023  
PUBLISHED: 12 May 2023

CITATION  
Ejaz U, Khan SM, Khalid N, Ahmad Z,  
Jehangir S, Fatima Rizvi Z, Lho LH, Han H  
and Raposo A (2023) Detoxifying the  
heavy metals: a multipronged study  
of tolerance strategies against  
heavy metal toxicity in plants.  
Front. Plant Sci. 14:1154571.  
doi: 10.3389/fpls.2023.1154571

COPYRIGHT  
© 2023 Ejaz, Khan, Khalid, Ahmad, Jehangir,  
Fatima Rizvi, Lho, Han and Raposo. This is an  
open-access article distributed under the  
terms of the [Creative Commons Attribution  
License \(CC BY\)](#). The use, distribution or  
reproduction in other forums is permitted,  
provided the original author(s) and the  
copyright owner(s) are credited and that  
the original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use, distribution or  
reproduction is permitted which does not  
comply with these terms.

## Detoxifying the heavy metals: a multipronged study of tolerance strategies against heavy metals toxicity in plants

Ujala Ejaz<sup>1</sup>, Shujaul Mulk Khan<sup>1,2\*</sup>, Noreen Khalid<sup>1</sup>,  
Zeeshan Ahmad<sup>1</sup>, Sadia Jehangir<sup>1</sup>, Zarrin Fatima Rizvi<sup>1</sup>,  
Linda Heejung Lho<sup>3\*</sup>, Heesup Han<sup>3\*</sup> and António Raposo<sup>4\*</sup>

<sup>1</sup>Department of Plant Sciences, Quaid-i-Azam University, Islamabad, Pakistan, <sup>2</sup>Ministry of Science and  
Technology, National University of Science and Technology (NUST), Islamabad, Pakistan, <sup>3</sup>Department of Plant  
Sciences, Seoul National University, College of Agriculture, Daegu 700-747, South Korea, <sup>4</sup>Department of  
Plant Sciences, University of Coimbra, Portugal, <sup>5</sup>Department of Plant Sciences, University of Coimbra,  
Portugal, <sup>6</sup>Department of Plant Sciences, University of Coimbra, Portugal, <sup>7</sup>Department of Plant Sciences,  
University of Coimbra, Portugal, <sup>8</sup>Department of Plant Sciences, University of Coimbra, Portugal, <sup>9</sup>Department of  
Plant Sciences, University of Coimbra, Portugal, <sup>10</sup>Department of Plant Sciences, University of Coimbra, Portugal

Heavy metal concentrations exceeding permissible limits threaten human life, plant life, and all other life forms. Different natural and anthropogenic activities emit toxic heavy metals in the soil, air, and water. Plants consume toxic heavy metals from their roots and foliar part inside the plant. Heavy metals may interfere with various aspects of the plants, such as biochemistry, bio-molecules, and physiological processes, which usually translate into morphological and anatomical changes. They use various strategies to deal with the toxic effects of heavy metal contamination. Some of these strategies include restricting heavy metals to the cell wall, vascular sequestration, and synthesis of various biochemical compounds, such as phyto-chelators and organic acids, to bind the free moving heavy metal ions so that the toxic effects are minimized. This review focuses on several aspects of genetics, molecular, and cell signaling levels, which integrate to produce a coordinated response to heavy metal toxicity and interpret the exact strategies behind the tolerance of heavy metals stress. It is suggested that various aspects of some model plant species must be thoroughly studied to comprehend the approaches of heavy metal tolerance to put that knowledge into practical use.

### KEYWORDS

heavy metal stress, plant defense mechanisms, genomics and transcriptomics, cell signaling pathways, plant structural and functional biology



# **Water Pollution Assessment in Aik-Stream: Leveraging Advanced Analytical and Machine Learning Techniques**

## **2.1 Introduction**

Water pollution primarily results from improper disposal of industrial, municipal, and domestic waste into water bodies (Chaudhry & Malik, 2017). Globally, an astonishing 2 million tons of sewage and various effluents are released into water systems daily, with developing nations facing an even graver scenario, where over 90% of raw sewage and 70% of untreated industrial waste finds its way into surface water sources (Azizullah, Khattak, Richter, & Häder, 2011). The significant sources of water pollution encompass domestic sewage, industrialization, population growth, pesticide fertilizer use, and the proliferation of plastics and polythene bags due to urbanization (Puckett, 1995).

### ***2.1.1 Water pollution by industries***

Water pollution from industrial discharges is widespread globally, particularly in rapidly developing countries where numerous factories release waste into local rivers. These factories generate a significant amount of wastewater, which contaminates streams and rivers. This wastewater often contains toxic and hazardous compounds, causing severe pollution of water ecosystems (Ado et al., 2015). Over the past few decades, industrial production of chemicals has been rising substantially. Notably, the effluents from some industries, such as tanneries, metalworking, and the apparel industry, are highly hazardous and harm the ecosystems and stream water quality they support (Zamora-Ledezma et al., 2021). According to projections, industrial wastewater production will more than double by 2025, causing freshwater sources to become more contaminated (Magdeline Hutton & Shafahi, 2019). This trend creates serious health hazards because people, animals, and plants can absorb metals from these pollutants. As a result, the brain system, kidney function, and other organs may be negatively impacted. According to the Lancet Commission on Pollution and Health Report 2018, pollution significantly contributes to more than 9 million preventable deaths yearly (P. J. Landrigan et al., 2019). People who depend on surface water and related water supply systems are at serious risk because of the widespread water pollution caused by the release of untreated industrial effluents. The Sustainable Development Goal (SDG) 6.3 of the UN intends to enhance water quality by 2030 by lowering pollution, limiting the release of



potentially harmful chemicals, reducing the amount of untreated wastewater, and significantly increasing recycling and safe reuse (P. J. Landrigan et al., 2019).

### ***2.1.2 The Health Impacts of Water Pollution***

The improper disposal of industrial waste and heavy metals can accumulate hazardous compounds in rivers and lakes, threatening people and wildlife. These pollutants cause immune system suppression, reproductive issues, and acute poisoning while also causing the spread of waterborne diseases like cholera and typhoid. Around 2.3 billion people are impacted by this water contamination problem, which substantially contributes to the world's health issues (Unesco, 2020). Due to dirty water and inadequate sanitation, developing countries face the brunt of the problem, where over 2.2 million people die each year (Organization, 2023). Approximately sixty percent of infant deaths are caused by water-borne infectious diseases, which are still a serious risk (Rashid et al., 2022). Several studies have also demonstrated a relationship between acute waterborne infections such as typhoid, hepatitis, cholera, dysentery, cryptosporidiosis, diarrhea, and water pollution (Fazal-ur-Rehman, 2019). Additionally, increasing water pollution levels raise the danger of developing cancerous conditions and "cancer villages" (Ayesha Tariq & Mushtaq, 2023). Notably, these health risks are more severe in emerging nations, which account for a significant portion of the 2.3 billion people suffering from water-related diseases worldwide, with countries like India and Pakistan experiencing considerable issues (B. S. Lal, 2019; Szálkai, 2023). However, these findings may not apply universally if there is a nonlinear relationship between pollution levels and health outcomes.

### ***2.1.3 Rising Concerns of Industrial Water Pollution in Pakistan***

Out of 6,634 registered industries in Pakistan, approximately 1,228 have been classified as highly polluting units (Sial, Chaudhary, Abbas, Latif, & Khan, 2006). This distinction arises from their significant contribution to water pollution, stemming from the discharge of organic and toxic substances in their waste effluents (Nasrullah, Bibi, Iqbal, & Durrani, 2006). Prominently featured among the culprits responsible for water pollution in Pakistan are industries spanning textiles, pharmaceuticals, ceramics, petrochemicals, food processing, steel production, oil mills, sugar refineries, fertilizer manufacturing, and leather tanning (Sial et al., 2006). These industrial sectors collectively release substantial volumes of wastewater into the environment, laden with various contaminants, including nitrates, nitrites, a spectrum of cations such as ( $\text{Ag}^+$ ,  $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Mg}^{2+}$ ,  $\text{Ca}^{2+}$ ), anions ( $\text{Cl}^-$ ,  $\text{CO}_3^{2-}$ ,  $\text{HCO}_3^-$ ), and noxious metals like arsenic, iron, lead, mercury, chromium, cadmium, copper, nickel, zinc,



cobalt, and magnesium (Davis III & Jacknow, 1975; R. Shrestha et al., 2021). These industries are most notably clustered in and around major urban centers. Their wastewater, rather indiscriminately, finds its way into nearby drains, rivers, streams, ponds, ditches, and even open or agricultural land (Sonone, Jadhav, Sankhla, & Kumar, 2020). An illustrative example lies in the River Kabul in Khyber Pakhtunkhwa, which bears the brunt of an estimated daily inflow of roughly 80,000 cubic meters (equivalent to  $8 \times 10^7$  liters) of industrial effluents (Siraj et al., 2022). Similarly, even the capital city of Islamabad grapples with the challenge of inadequate effluent management within its two industrial estates, leading to the unbridled release of waste directly into the Soan River (Zakaullah & Ejaz, 2020).

Perhaps the most disconcerting aspect of this scenario is the stark statistic that merely 1% of the wastewater generated by industries in Pakistan undergoes any form of treatment before being unceremoniously discharged (Mulk, Korai, Azizullah, & Khattak, 2016). Consequently, wastewater laden with potentially hazardous substances is funneled into water bodies, all without due regard for the environmental perils it poses. The cumulative magnitude of this issue is staggering, with an estimated 40 billion liters of waste effluent finding its way into Pakistan's water bodies each day, courtesy of various industries (Henneberry, Khan, & Piewthongngam, 2000).

#### ***2.1.4 Industrial Pollution in Sialkot, Pakistan***

Sialkot is renowned as the epicenter of export-oriented industries in Pakistan, particularly for its prominent role in manufacturing surgical instruments, sports equipment, and leather goods. The city hosts 3,229 industries, comprising 270 leather tanneries, 220 surgical equipment manufacturers, and 900 sports goods factories (Abdul Qadir, Riffat Naseem Malik, & Syed Z Husain, 2008a). Amongst these manufacturing, leather tanneries stand out as the leading contributors to the generation of sewage sludge and wastewater. These leather tanneries manufacture millions of leather items daily and discharge sizeable volumes of untreated wastewater into adjacent streams. The Cleaner Production Center (CPC) calculated that Sialkot produces 297 tons of leather daily, with tanneries releasing 215,036.1 gallons of effluents daily (A. Qadir et al., 2008a). The quality of the water is seriously threatened by this discharge, including activities like bathing, stripping, tanning, dyeing, and disposal of grease liquor. A significant source of contamination for the region's soil and water systems is the effluent from the leather tanneries in Sialkot, according to empirical studies by (Malik, Jadoon, et al., 2010; A. Qadir et al., 2008a). These effluent also



contain metals, the main one being chromium, as well as sodium, calcium, organic solvents, and suspended particles. These discharges also contain a variety of contaminants and salts (Chowdhury, Mostafa, Biswas, Mandal, & Saha, 2015). These effluents from leather tanneries are observed as the leading environmental pollutants owing to their makeup, which contains more than 40 different chemicals, such as heavy metals, acids, and dyes (Chowdhury, Mostafa, Biswas, & Saha, 2013; Khalid et al., 2021b). This waste encompasses used animal hides and effluent laden with various chemicals, including chlorides, sulfates, hydrocarbons, amines, aldehydes, and heavy metals like arsenic and chromium used in the leather tanning process (Bosnic, Buljan, & Daniels, 2000). Previous studies have highlighted the health hazards of these industries to Sialkot's local human population and wildlife (Junaid et al., 2016; Junaid, Hashmi, Tang, Malik, & Pei, 2017).

### ***3.1.5 The Aik-Stream of Sialkot, Pakistan***

The Aik-stream, a significant Indo-Pak transboundary tributary of the Chenab River flowing through Sialkot, Pakistan, is paramount in the regional hydrological context. Nevertheless, its ecological importance has been compromised due to a pronounced environmental challenge. As it courses through the urban landscape, this waterway receives a substantial influx of polluted wastewater. This effluent contains harmful constituents, including toxic chemicals and heavy metals, primarily from municipal waste and industrial discharge sources. Unfortunately, this contaminated flow is directed straight into the Chenab River without prior treatment or remediation. Remarkably, the Aik-stream absorbs an astonishing daily load of 52 million liters of wastewater, further exacerbated by an additional 1.1 million units of waste discharged by nearby leather-producing factories (A. Qadir & Malik, 2009). Wastewater discharge contains an amalgamation of organic and inorganic compounds encompassing various hazardous substances. This includes heavy metals, synthetically produced tannins, oils, resins, bio-toxins, and disinfectants, each contributing to a complex and potentially toxic mix. These substances pose significant health risks, impacting the environment and public health (Tadesse, Guya, & Walabu, 2017).

Such a varied and dangerous cocktail of contaminants necessitates urgent attention and effective management strategies to mitigate the adverse health effects associated with exposure to these pollutants (Bosnic et al., 2000). The consequences of this ecological degradation are profound. Empirical research consistently highlights the presence of alarmingly elevated levels of heavy metals, including chromium, lead, cadmium, mercury, copper, zinc, nickel, and arsenic, within the compromised water of the Aik-stream (Butt et al.,



2021; Khalid et al., 2021b; Murtazaa & Usmanb, 2022). In its historical records, the Aik Stream was renowned for its pristine waters, once serving as a source for domestic water needs (A. Qadir & Malik, 2009). However, this idyllic past has been replaced by an urgent environmental concern. It is crucial to emphasize that the surface water above the Water from this stream is currently used directly for irrigating food crops and fodder, introducing contaminants into the entire food chain. The direct discharge of industrial and municipal wastewater into these aquatic channels is a conspicuous and prominent contributor to this alarming contamination scenario.

### ***2.1.6 Machine learning as a novel monitoring tool***

The transformation of Aik-stream into open sewers is a direct consequence of effluents from heavily industrialized and densely populated regions. To manage surface water quality effectively, it is imperative to maintain ongoing surveillance. This is vital for safeguarding ecosystems, ensuring public health, managing water resources sustainably, controlling pollution, and guiding policy formulation (Brack et al., 2017; Geissen et al., 2015). Conducting water quality monitoring and assessment is crucial for identifying potential environmental and health risks, safeguarding natural ecosystems, and ensuring the availability of clean water resources for both present and future populations. However, the reliance on conventional monitoring methodologies alone is insufficient. The incorporation of advanced predictive models or mechanisms is imperative. These tools can rapidly forecast potential risks, offering a proactive approach to mitigating adverse impacts on water quality and aligning with sustainable water management and environmental protection objectives (A. R. M. T. Islam et al., 2021). Building on these insights, a range of proactive measures can be adopted. These include initiating preemptive treatments to neutralize contaminants, diversifying or altering water sources, issuing prompt public advisories to ensure community safety, modifying and upgrading water treatment infrastructure, and developing emergency response strategies to address identified threats. Such measures are essential for maintaining water quality and safeguarding public health while enhancing water management systems' resilience against potential environmental risks (Berglund et al., 2020; Sun & Scanlon, 2019). Forecasting water quality is critically important for several socio-economic sectors reliant on clean and safe water supplies. With technological advancements, artificial intelligence (AI) has emerged as a critical player in developing sophisticated algorithms and predictive techniques. These AI-driven methods enable a more detailed and comprehensive analysis of data related to water quality. The current research explores the primary methodologies



employed in AI-based water quality prediction, explicitly examining the capabilities of two machine learning algorithms: Gradient Boosting (GB) and Random Forests (RF). This research aims to improve the accuracy of water quality predictions while addressing the complexities inherent in the datasets used. These models, GB and RF, are particularly adept at predicting water quality thanks to their ability to handle various monitoring data, manage complex data interrelations, adapt to different types of data, and accurately quantify the uncertainty in predictions, which is crucial for effective water resource management. Ensemble methods, such as Gradient Boosting and Random Forests, stand out for their proficiency in capturing intricate data relationships. They do this by amalgamating predictions from many models, thereby significantly boosting performance, especially in scenarios characterized by complex, non-linear data patterns. This research will delve into these models, unraveling their potential to enhance the predictive accuracy of water quality assessments and their implications for water resource management. (Mienye & Sun, 2022).

#### ***2.1.7 Problem and Novelty Statement***

The Aik-Stream, a significant natural tributary of the River Chenab, flows through Sialkot, with over 2.5 million people in its catchment area (A. Qadir & Malik, 2009). Historically, this stream was a crucial resource for drinking water, domestic usage, and irrigation. The land around the stream, characterized by alluvial deposits due to regular flooding, has been under cultivation since prehistoric times. However, the Aik-Stream faces significant environmental challenges as it receives a substantial volume of effluents, untreated sewage, and solid waste (M. S. Khan et al., 2019). These pollutants significantly alter the stream's physical, chemical, and biological properties. Most industrial activities are in urban areas along major roads and highways. Much of the effluents come from tanneries, which generate considerable liquid waste containing around 130 chemicals and municipal sewage (A. Qadir & Malik, 2009). No facilities have been established to treat industrial effluents or sewage, leaving untreated wastewater from these sources. This absence of treatment infrastructure contributes to the escalating pollution levels in the Aik-Stream, raising concerns about environmental and public health in the region.

This research project was initiated to address the environmental challenges of the Aik-stream. To achieve this, accurate and statistically valid data was essential to demonstrate the impact of industrial activities on the stream's water quality. There is a critical need for an in-depth analysis to identify the patterns and changes in the water quality of these streams and to understand the implications of current and future contamination risks. Consistent monitoring



is crucial for effectively managing surface water quality and is critical for protecting ecosystems, safeguarding human health, ensuring sustainable management of water resources, controlling pollution, and informing policy-making (Brack et al., 2017). However, reliance on traditional monitoring techniques alone is insufficient. Adopting advanced methods or models that can quickly identify potential risks is often necessary, thus preventing adverse effects on water quality (A. R. M. T. Islam et al., 2021).

This study also seeks to provide a comprehensive overview of the principal methods used in AI-based water quality prediction, which is intended to improve the monitoring process. These models can efficiently predict water quality with reduced computational time and lower costs, and they simplify the broad array of parameters required for water quality assessment and monitoring into a more manageable subset. No effort has been made to evaluate these innovative monitoring tools in any polluted water resource in Pakistan. The findings of this research are intended to serve as a guide for sustainable watershed management, intelligent urban planning, environmentally friendly industrial practices, and sustainable agriculture, not only in the Aik-stream catchment area but also in similar ecosystems.

#### ***2.1.8 Hypothesis:***

1. The levels of heavy metals and critical physico-chemical parameters predominantly influence the spatial variations in Aik-stream's water quality.
2. Advanced AI models will substantially improve the prediction and understanding of these spatial variations in water quality, offering a robust tool for assessment and management in Aik-stream.

The main objectives of this chapter are to:

#### ***2.1.9 Objectives:***

1. Examine spatial differences in Aik-stream's water quality by identifying key water quality parameters.
2. Determine the origins of heavy metals and physicochemical elements in surface water, emphasizing variables significantly influencing water quality variations.
3. Develop and compare various artificial intelligence models for predicting surface water WQI, identifying the most efficient models for this task.



## 2.2 Materials and Methods

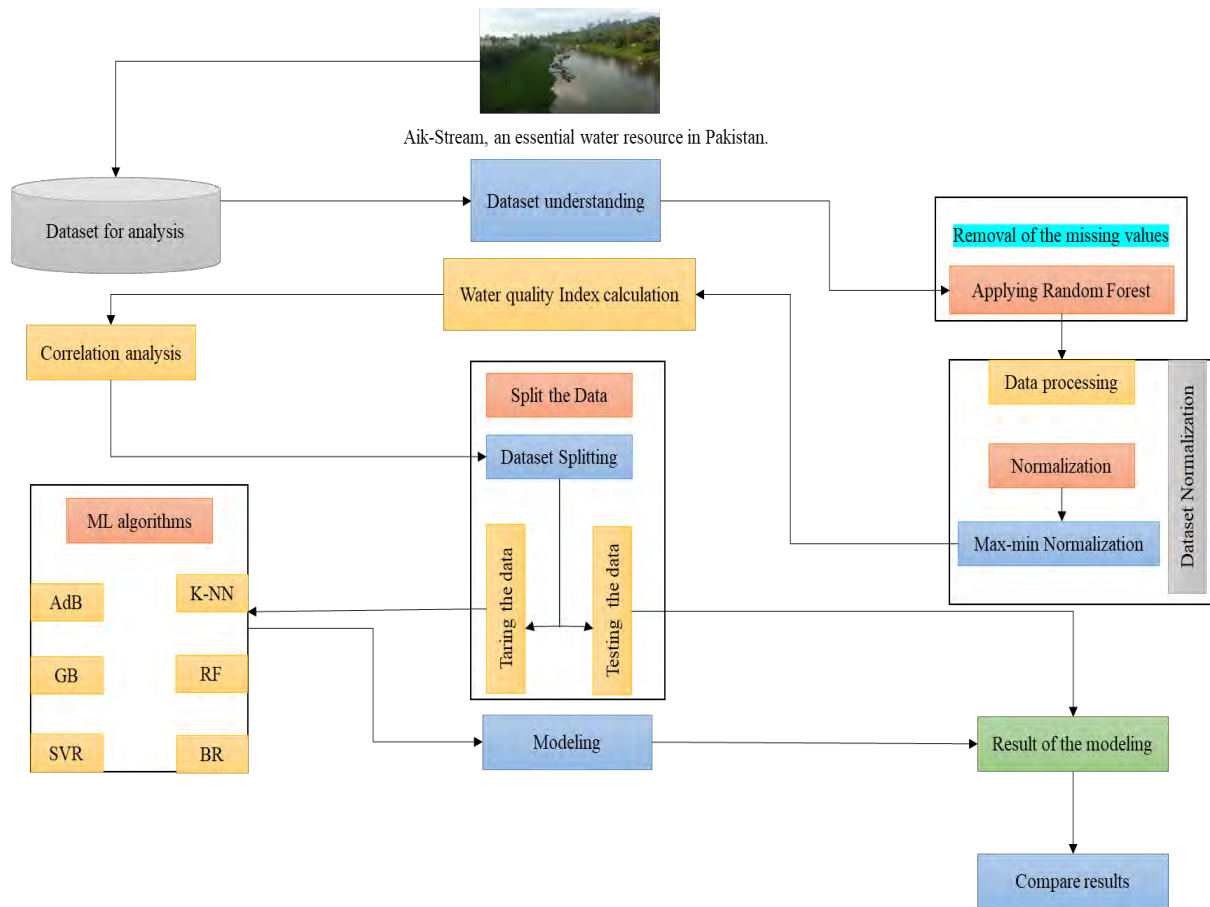


Figure 2. 1 Propose model working diagram

### 2.2.1 Study Area

The Aik-Stream (32°63'N- 74°99'E and 32°45'N- 74°69'E), a significant tributary of the Chenab River flowing through the city of Sialkot, Pakistan, is the main focus of this paper. (Fig. 2.2). Sialkot is in a humid subtropical climatic zone, where the average annual precipitation measures 957.9 mm, equivalent to 37.7 inches (S. Ali, Zhang, & Yue, 2020). The stream starts from the Pir Punjal Range in the Lesser Himalayan region of Jammu and Kashmir at 530 meters at sea level (Malik, Jadoon, et al., 2010) (Fig 2.2). It spans 131.6 km, with a 315 Cs annual flow rate and a catchment area covering approximately 1,062 km<sup>2</sup>. Typically, the stream receives its lowest discharge during the early summer, while the utmost occurs during the monsoon season (Mahmood et al., 2014). As the stream flows through the city, it receives a substantial influx of wastewater, including toxic chemicals and heavy metals from municipal waste and industries, which drain into the Chenab River without treatment. This stream receives 52 million liters of wastewater daily, with an additional 1.1



million waste units from leather-producing factories (Daily, 2006; Pakistan, 2007; A. Qadir & Malik, 2009; A. Qadir et al., 2008b). The leather industry waste discharge includes organic and inorganic substances, toxic materials, i.e., heavy metals, chemically synthesized tannins, oils, resins, bio-toxins, and disinfectants (Garai, 2014; Maqbool et al., 2018; Rabelo et al., 2018; S. R. Tariq et al., 2010). Researchers have reported high levels of heavy metals such as chromium, lead, cadmium, mercury, copper, zinc, nickel, and arsenic in the Sialkot Industrial Zone (Jadoon & Malik, 2019; Khalid et al., 2021a; Lokhande, Singare, & Pimple, 2011; Malik, Jadoon, et al., 2010; A. Qadir & Malik, 2009; A. Qadir et al., 2008a).

The Cleaner Production Centre (CPC) in Sialkot estimates that the ideal leather production level is approximately 297 tons daily, resulting in 9,388 cubic meters of tannery effluent daily. On average, each tannery in the Sialkot district generates between 547 and 814 cubic meters of effluent per day, culminating in 11,000 cubic meters (A. Qadir et al., 2008a). In contrast, units manufacturing surgical instruments produce effluents containing acids and dissolved metals from electroplating processes, but the volume of these effluents is significantly less than that from tanneries. These surgical instruments and sports goods contribute minimally to effluent production.

Another significant pollution source is municipal sewage. Sialkot City lacks a municipal waste treatment plant, leading to untreated sewage, encompassing organic and inorganic pollutants and high levels of suspended organic material, along with animal and human waste, being discharged directly or indirectly into local water bodies. Annually, the city generates about 19 million cubic meters of wastewater (Bhatti & Latif, 2011), disposed of in natural streams, ponds, and agricultural lands.



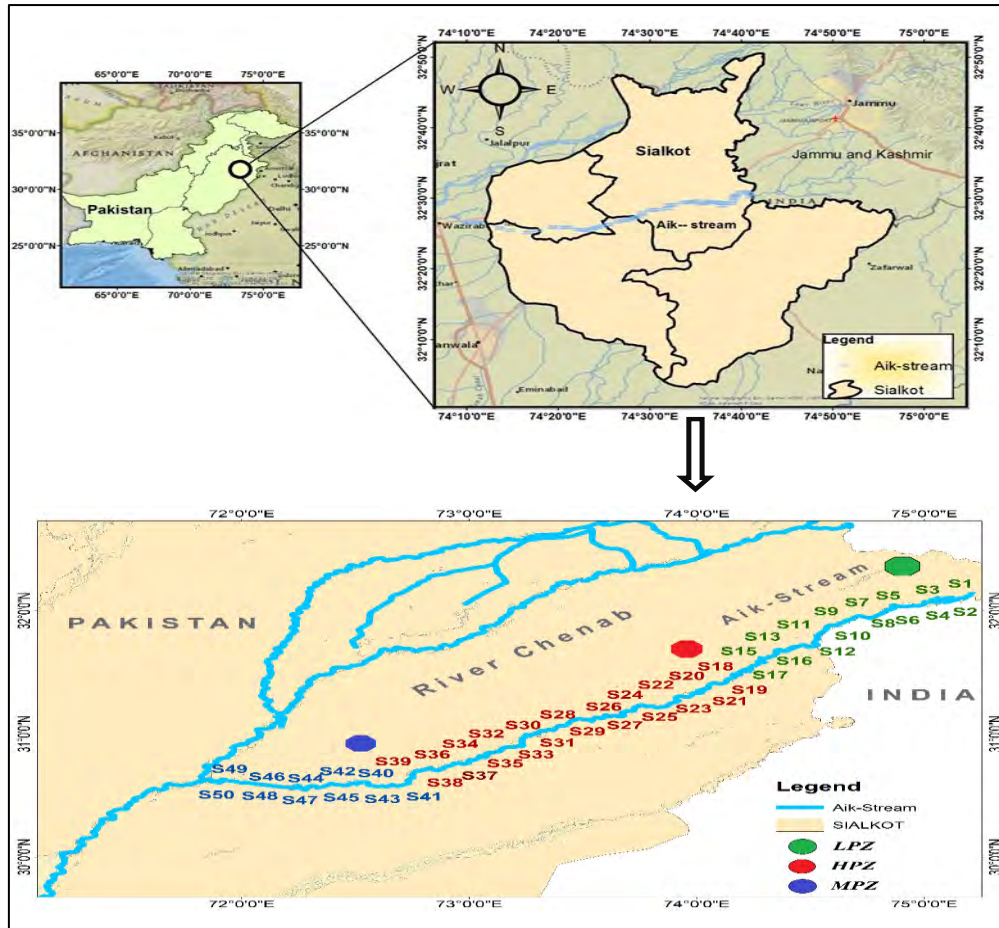


Figure 2. 2 Map depicting the industrialized city of Sialkot in Pakistan, along the Aik-stream with distinct polluted zones i.e. Less Polluted Zone (LPZ), High Polluted Zone (HPZ) and Moderately Polluted Zone (MPZ) (Source: Arc map 10.5).

### 2.2.2 Study Design

This study aimed to assess the water quality of Aik Stream, particularly concerning the impact of industrial wastewater. To measure the water's organic, physical, and physiochemical attributes, the trap research methodology collected data from multiple points along the stream, including upstream, midstream, and downstream locations. In conjunction with this quantitative analysis, a qualitative approach was adopted to identify and understand the varying pollution levels at different stream sections. This blend of quantitative and qualitative data was pivotal in providing a comprehensive view of the stream's condition, allowing for a detailed examination of how industrial wastewater influences the water quality of Aik Stream and its potential risks to the ecosystem. This holistic approach not only elucidated the current state of the stream but also highlighted the environmental implications of industrial waste discharge, offering valuable insights for future conservation and management strategies.



### **2.2.3 Data collection**

The water samples were collected at 150 sampling sites (150×3 samples) from the Aik-stream at one-kilometer intervals, ensuring comprehensive stream flow coverage from its source to the endpoint. Trap water techniques were implemented to collect water samples in April and the end of September, corresponding to the hydrologically high and low flow periods. Water samples were taken 30cm below the water surface, within the 100m range of a sampling site, halfway from the water column in a direction opposite to the flow. Water samples were collected 30cm beneath the surface, within a 100m radius of each sampling site, and were taken from halfway up the water column against the current's flow. The sampling container was carefully filled with water in the sampling bottles, ensuring no air bubbles were trapped. Three-stream water sub-samples were mixed at each site to form a composite sample. These samples were stored in plastic bottles previously cleaned with HNO<sub>3</sub> and were securely sealed to prevent any leakage during transportation. To preserve their integrity, all water samples were placed in ice boxes and transported to the laboratory following the standard method outlined by (Rodier et al., 2009). The samples were examined to determine various essential characteristics related to physicochemical and biochemical parameters.

A brief description of each site of the Aik-stream is given below:

### **2.2.4 Sampling sites (S1-S51)**

Sampling sites (S1-S51) were chosen along the upper section of the Aik-Stream, beginning from the boundary between Pakistan and the Indian-controlled Jammu & Kashmir, where the Aik-Stream flows into Pakistan, and terminating before it reaches Sialkot city (Fig. 2.3 (a-d)). These sites are located upstream of the Aik-Stream and are relatively undisturbed with commendable water quality. Several water pumps are positioned along the stream's banks for irrigation purposes. While no direct sources of pollution were detected, agricultural runoff and atmospheric deposits from the Jammu district, under Indian control, might contribute as indirect pollution sources.



a



b





c



d



Figure 2. 3 (a-d) illustrates the upstream area of the Aik-Stream, showcasing its relatively undisturbed nature and good water quality.

### ***2.2.5 Sampling sites (S51-S117)***



Sampling locations (S51-S117) were in the midstream region of Aik-Stream, starting from Sialkot city and ending at Sambryal town. This section is exposed to effluents from small industries, tanneries, and municipal waste from Sialkot city. Many of these effluents

a



originate from the industrial sector along this midstream stretch (Fig. 2.4 (a-d)).





b



c





d



Figure 2. 4 (a-d) illustrates the mid-stream area of the Aik-Stream, showcasing its relatively disturbed nature and impaired water quality.

#### **2.2.6 Sampling sites (S117-150)**

Sampling sites (S117-S150) were situated approximately 24km downstream from Sialkot, near Wazirabad city. The final sites were positioned just before the water merged into the Palkhu stream, which eventually flows into the Chenab River (Fig. 2.5 (a-b)). Effluents from tanneries along the Sialkot-Wazirabad route, surface runoff from agricultural activities, and domestic sewage from surrounding towns are discharged into this site.

a





b



Figure 2. 5 (a-b) illustrates the downstream area of the Aik-Stream, showcasing its moderately disturbed nature and marginal water quality

#### ***2.2.7 Analytical Procedures in the Laboratory***

All samples were transported to the Plant Ecology Laboratory at Quaid-i-Azam University Islamabad, where they were immediately stored in a freezer to prevent any potential physicochemical alterations. Before conducting metal analyses, all other water quality tests were performed. The water samples designated for dissolved metal content and particulate matter assessments were separated. To ensure that metals remained dissolved during storage, the samples for dissolved metals were acidified with a few drops of  $\text{HNO}_3$ . In the laboratory, a range of water quality parameters were determined, including pH, Temperature ( $T$ ,  $^{\circ}\text{C}$ ), Electrical Conductivity (EC,  $\mu\text{S}/\text{cm}$ ), Total Dissolved Solids (TDS,  $\text{mg}/\text{L}$ ), Total Suspended Solids (TSS,  $\text{mg}/\text{L}$ ), Biological Oxygen Demand (BOD,  $\text{mg}/\text{L}$ ), Chemical Oxygen Demand (COD,  $\text{mg}/\text{L}$ ), Ammonia Nitrogen ( $\text{NH}_3\text{-N}$ ,  $\text{mg}/\text{L}$ ), Total Organic Carbon (TOC,  $\text{mg}/\text{L}$ ), Chloride ( $\text{Cl}^-$ ,  $\text{mg}/\text{L}$ ), Nitrate Nitrogen ( $\text{NO}_3\text{-N}$ ,  $\text{mg}/\text{L}$ ), and Oil and Grease (O&G,  $\text{mg}/\text{L}$ ). Additionally, the levels of various metals, such as zinc (Zn), nickel (Ni), copper (Cu), chromium (Cr), lead (Pb), arsenic (As), and mercury (Hg) were also determined, as detailed in Table 2.1. Some glimpses of the laboratory work can be seen in Fig. 2.6 (a-j).



The physiochemical parameters of the water are determined through McLean methods (McLean 1982). The TDS and pH were measured using TDS meters and pH (Russel RL060P). Heavy metals (Zn, Cu, Ni, Cr, As, Cd, Pb, and Hg ) were observed by atomic absorption spectrophotometers VARIAN, AA240FS, using the approach described by (Cantle, 1986). For metal analysis, acid digestion was carried out using a 250 mL flask containing 150 mL of the wastewater sample. To this sample, 10 mL of a mixed solution of Perchloric Acid ( $\text{HClO}_4$ ) and Nitric Acid ( $\text{HNO}_3$ ) in a 1:3 ratio was added. The flask was then allowed to rest for 24 hours. The digestion process involved heating the sample on a hot plate, starting at  $150^\circ\text{C}$  for the first hour, then increasing to  $235^\circ\text{C}$ . This heating continued until the nitric acid's red fumes ceased, and white fumes appeared. After cooling, the digested sample was filtered using No. 42 filter paper. The resulting filtrate was then subjected to analysis through atomic absorption spectrophotometry.

Assessing organic parameters in water plays a vital role in environmental monitoring, identification, and treatment of wastewater. Key organic characteristics of water are typically evaluated using tests for Biochemical Oxygen Demand (BOD), Chemical Oxygen Demand (COD), Total Organic Carbon (TOC), and Oil and grease (O&G) as outlined by APHA (2005). BOD is a traditional method for determining the amount of organic matter in water samples. The process involves adjusting the pH, removing chlorine from the samples, and then incubating them in special biochemical oxygen demand bottles for five days. The oxygen levels are measured at the start and end of this period, and the difference, attributed to microbial activity, indicates the BOD concentration in mg/L. COD is a quick and effective alternative to BOD for assessing organic matter concentration. It employs chemical oxidation to quantify the concentration of organic and inorganic compounds. This method involves setting up standard solutions, adding sample aliquots to pre-prepared bottles, and inducing a controlled chemical reaction. The resultant colour change is measured using a colourimeter, from which the COD concentration is calculated. This method is beneficial for wastewater treatment personnel to adjust treatment processes in response to varying wastewater compositions.

TOC analysis is gaining traction for its rapid execution, taking just 5-10 minutes to complete. It measures the total organic carbon in water samples and can estimate BOD levels once a consistent TOC to BOD ratio is established. Various techniques involving heat and oxygen, ultraviolet radiation, or chemical oxidants are used in carbon analysis instruments, typically followed by infrared analysis to convert organic carbon to  $\text{CO}_2$ . TOC testing is



valued for its quick results and informative insights into the water's organic matter content. The O&G in wastewater, originating from sources like plant and animal fats or petroleum products, poses challenges due to their hydrophobic nature and tendency to form emulsions. For the O&G test, water samples are collected in glass containers, acidified, and then extracted using a solvent such as n-hexane. Post-extraction, the solvent is distilled off to isolate the O&G, and their concentration is calculated based on their weight and the initial sample volume. This test is essential in determining the presence of oils and grease in wastewater, impacting treatment processes and wastewater system management. For a detailed analysis of these water quality parameters, refer to Table 2.1.

Table 2. 1 Description and measurement techniques of the explanatory variables.

Variables	Description	Measurement techniques
pH	The pH of water measures the acidity or alkalinity.	pH meter was used to determine the pH value by measuring the hydrogen ions in water.
T (°C)	A water sample's temperature indicates its molecules' average kinetic energy.	A calibrated thermometer is submerged in the water sample until stability is reached. The temperature is then read and recorded.
COD (mg/L)	The chemical oxygen demand measures the number of organic pollutants present in water.	Potassium dichromate oxidizes organic matter in the water sample to determine COD, and the resulting color change is quantified using a spectrophotometer to assess the concentration of organic content.
BOD (mg/L)	Biochemical oxygen demand measures how much organic matter microorganisms can decompose in water.	The water sample's dissolved oxygen (DO) concentration is determined using a meter first. A nutrient-rich seed solution is added, followed by a 5-day incubation, after which the final DO is measured. BOD is calculated as (Initial DO - Final DO) x (Dilution Factor).
TDS (mg/L)	Total dissolved solids refer to the concentration of the inorganic and organic substances dissolved in water.	A calibrated TDS meter measured the concentration of inorganic and organic compounds dissolved in water.
TSS (mg/L)	Total suspended solid measures the concentration of the solid particles suspended in water.	The water was filtered through a pre-weighed filter, which was then dried in an oven and weighed to capture and determine the weight of the solids for the quantification of TSS in water.
NH <sub>3</sub> -N (mg/L)	Ammonia nitrogen is a measure of the concentration of ammonia in water.	The concentration of NH <sub>3</sub> -N in water is tested by adding sulfuric acid to stabilize the ammonia concentration; the colorimetric approach is then utilized to accurately detect ammonia levels in water samples.
O& G (mg/L)	Oil and grease measure the concentration of oil and grease contaminants in water.	To determine oil and grease (O&G) in water, the sample is extracted with cyclohexane, dried using anhydrous sodium sulfate, and then weighed. The concentration is calculated by dividing the extract's weight by the water sample's volume.
TOC (mg/L)	Total organic carbon measures the concentration of all organic carbon compounds in water.	TOC in water is assessed by incinerating the sample in a high-temperature furnace, where organic carbon is converted into CO <sub>2</sub> . This CO <sub>2</sub> is then detected using the UV-persulfate oxidation method, and the TOC concentration



		in the original sample is calculated using known reaction stoichiometry.
NO <sub>3</sub> -N (mg/L)	Nitrate nitrogen measures the concentration of nitrate ions in water.	Nitrate-nitrogen (NO <sub>3</sub> -N) in water is determined by chilling a sample before forming a colored complex with nitrate ions in automated analyzers. This color's intensity is then measured and converted into NO <sub>3</sub> -N concentration.
Cl <sup>-</sup> (mg/L)	Chloride indicates the concentration of chloride ions in water.	Titration with silver nitrate (AgNO <sub>3</sub> ) solution is used to determine the Cl <sup>-</sup> concentration of water. As silver ions react with chloride ions, a white precipitate of silver chloride (AgCl) forms and the amount of AgNO <sub>3</sub> needed shows the chloride concentration.
Cr (mg/L)	Chromium is a measure of the concentration of chromium ions in water.	An atomic absorption spectrophotometer determines the amount of chromium (Cr) in water. This approach includes sending a laser beam through a water sample and measuring the absorption of various wavelengths of light by chromium atoms, allowing the chromium concentration in the water to be determined.
Zn (mg/L)	The concentration of the Zinc ions in water.	An atomic absorption spectrophotometer is used to measure zinc (Zn) in water and analyze the zinc content.
Cu (mg/L)	Copper is measured by the concentration of the copper ions in water.	An atomic absorption spectrophotometer assessed the copper (Cu) content in water.
Pb (mg/L)	Lead is the concentration of lead ions in water.	An atomic absorption spectrophotometer was used to measure lead (Pb) in water to analyze the lead content.
Cd (mg/L)	Cadmium is a measure of the concentration of cadmium ions in water.	The cadmium (Cd) content in water was analyzed using an atomic absorption spectrophotometer to measure its concentration.
Ni (mg/L)	Nickel is a measure of the concentration of nickel ions in water.	An atomic absorption spectrophotometer was used to measure nickel (Ni) in water to analyze the nickel content.
As (mg/L)	Argon showed the concentration of argon ions in water.	An atomic absorption spectrophotometer was used to measure Arsenic (As) in water to analyze the arsenic content.
Hg (mg/L)	Mercury is a measure of the concentration of mercury ions in water.	An atomic absorption spectrophotometer was utilized to analyze the mercury (Hg) content in water for measurement.





Figure 2. 6 Some glimpse of the laboratory work.

### 2.2.8 Statistical analysis

Statistics analyses were performed using (R software version 4.0.2) and (python version 3.5) software. Three conventional multivariate techniques, such as Hierarchical Algorithmic Cluster Analysis (HACA), The Analysis of Variance (ANOVA), and Principal Component Analysis (PCA), were used for the water quality assessment and interpretation of the results. These multivariate statistical techniques have been widely used in various studies for the interpretation of chemical/physical characteristics of water quality parameters (Jahin, Abuzaid, & Abdellatif, 2020; K. R. Singh, Goswami, Kalamdhad, & Kumar, 2020)



#### *2.2.8.1 Hierarchical Agglomerative Cluster Analysis (HACA)*

The stream water quality data set was subjected to HACA to identify clusters of the sampling sites indicating their similarity based on water quality parameters (Table 2.1). For this purpose, 19 water quality parameters measured from 150 sites (19 variables  $\times$  150 cases) were subjected to HACA analysis. HACA was performed on the mean values of water quality parameters. Euclidean distances were chosen as a measure of linkage that uses analysis of variance to evaluate the distances between clusters, attempting to minimize the sum squares of any two clusters that can be formed at each step (Kavitha, Gayathri, & Devaraj, 2023). Eight metals were also subjected to HACA to determine the association of metals and their possible sources in water quality data. Pearson correlation was used to confirm the HACA results and find an association between different water quality parameters.

#### *2.2.8.2 Analysis of Variance (ANOVA)*

The Analysis of Variance (ANOVA) was applied to the Aik-stream's water quality dataset to ascertain if significant mean differences existed among various groups. ANOVA specifically evaluated the variances both between and within these groups. The null hypothesis asserted that there are no significant differences in the means of the water samples from the Aik-stream. Conversely, the alternative hypothesis suggested a significant mean difference between at least one set of groups. This analytical approach offers valuable insights into the disparities and trends in water quality at different sampling locations within the Aik-stream.

#### *2.2.8.3 Principal Component Analysis (PCA)*

The principal component analysis was used to identify the important physiochemical parameters that affect the water quality of the surface water and to investigate the possible source of different pollutants. Principle component analysis was performed on three zones of the study area: LPZ, MPZ, and HPZ. Principle Component Analysis (PCA) is a powerful pattern recognition technique that attempts to describe a large set of intercorrelated variables with a smaller set of independent variables Singh et al., (2005a). PCA begins by extracting the eigenvalues and eigenvectors from the covariance matrix, which describes the dispersion of the original values. Eigenvectors represent the list of values when initial correlated parameters are multiple by new uncorrected parameters of the PC.



### 2.2.9 Innovative Monitoring Approaches Using Machine Learning Models

In addition to these conventional techniques, we explored a novel approach by incorporating machine learning models to predict WQI in less time with fewer variables. These models aid in monitoring, assessing, and predicting water quality, offering a more efficient surveillance method.

### 2.2.10 The Water Quality Index (WQI)

Various methods and tools are utilized to assess water quality, and one prominent approach is the Water Quality Index (WQI). This method effectively summarizes extensive data on water quality into a single numerical index score, streamlining the evaluation process (Parween, Siddique, Diganta, Olbert, & Uddin, 2022). The Water Quality Index (WQI), designed by the Canadian Council of Ministers of the Environment (CCME) in 2001, was used in this study. It is categorized as an open index, meaning it does not predefine specific parameters or quality criteria for its calculation. Instead, the index requires using a set of parameters pertinent to the particular body of water being assessed and the intended purpose of the evaluation. Therefore, this index is a versatile tool that can be customized to suit various water quality monitoring programs. However, it is essential to note that meaningful comparisons of results can only be made when the same parameters and criteria are consistently applied.

The CCME-WQI is calculated using the three numerical factors of scope (F1), frequency (F2), and amplitude (F3). The fraction of parameters surpassing the quality standards at least once depends on the scope factor, represented by Equation 2.1. Equation 2.2 represents the frequency factor, which expresses the proportion of analytical results that exceed the criteria when all parameters are considered. The amplitude factor, represented by Equation 2.3, computes the difference between the quality criteria and the analytical results that do not meet the standards. Collectively, these characteristics enable the WQI to provide a comprehensive assessment of water quality.

$$\text{Scope: } F1 = \left( \frac{\text{Number of failed variables}}{\text{Total number of variables}} \right) \times 100 \dots \dots \dots (2.1)$$

$$\text{Frequency: } F2 = \left( \frac{\text{Number of failed tests}}{\text{Total number of tests}} \right) \times 100 \dots \dots \dots (2.2)$$

$$\text{Amplitude: } F3 = \frac{nse}{0.01nse + 0.01} \dots \dots \dots (2.3)$$



The calculation of Factor 3 (Amplitude) involves a three-step process, which is used to determine the difference between the desired value and each concentration. These steps are represented by equations 2.4, 2.5, and 2.6 in the amplitude formulation.

$$\text{Excursion}_i = \left( \frac{\text{Failed test value}_i}{\text{Objective}_j} \right) - 1 \dots \dots \dots (2.4)$$

The test value should not be less than the target value in the following circumstances:

$$\text{Excursion}_i = \left( \frac{\text{Objective}_j}{\text{Failed test value}_i} \right) - 1 \dots \dots \dots (2.5)$$

The total deviation coefficients of all individual results, no matter whether they meet the predetermined objectives, must be added up and divided by the sum of the number of individual results to determine the overall extent of non-compliance; this value is represented by the normalized total of all the deviation coefficients (NSE), which is written as:

$$\text{NSE} = \left( \frac{\sum_{i=1}^n \text{Excursion}_i}{\text{Number of tests}} \right) \dots \dots \dots (2.6)$$

NSE from the targets is then mapped using an asymptotic function, assigning values in the range of 0 to 100 to calculate term F3 (equations 2.6).

After determining these factors, the actual WQI can be computed by these three terms as vectors and incorporated together. The sum of each factor's squares equals the index's square (equations 2.7). The WQI is conceptualized in this approach as a three-dimensional space, with each axis representing one of the three factors. According to the CCME's definition, the index is directly related to these factors.

$$I = \left[ \frac{\sqrt{F_1^2 + F_2^2 + F_3^2}}{1.732} \right] \dots \dots \dots (2.7)$$

The calculated values are normalized using the divisor 1.732, so the Water Quality Index (WQI) result is from 0 to 100. The water quality is evaluated on this scale from 0 to 100, with 100 representing the highest quality. The CCME has established five categories based on water quality, as shown in Table 2.2.



Table 2. 2 Classification of CCME-WQI and Corresponding Water Status

<b>Water Quality Class</b>	<b>Index Range</b>	<b>Water Status</b>
Excellent	95-100	Very good quality water
Good	80-94	Good quality water
Fair	65-79	Acceptable quality water
Marginal	45-64	Poor quality water
Poor	0-44	Very poor-quality water

### ***2.2.11 Machine learning models***

Two machine learning models, Gradient Boost (GB) and Random Forest (RF), were chosen to be trained on the water quality dataset. Their performance indicators were considered in making this choice, and an assessment was conducted regarding how well they aligned with specific requirements. Gradient Boost was selected because it can effectively capture intricate interactions in the data (Z. Zhou et al., 2020). It is an ensemble method that builds models one at a time, correcting the flaws of the previous model as it goes. GB is skilled at handling heterogeneous data and generating highly accurate predictions (Polikar, 2012). However, if the learning rate is slow or the number of boosting iterations is excessive, it may be subject to overfitting. Hyperparameters must be tuned carefully (Uddin, Nash, Rahman, & Olbert, 2022). RF can manage interactions, outliers, and high-dimensional data. By averaging predictions from various trees, this collection of decision trees lessens overfitting. The robustness, computational effectiveness, and provision of feature importance measures of RF are well known. In contrast to other ensemble methods, however, RF can be more challenging to interpret than a single decision tree and may have trouble capturing complex nonlinear relationships.

#### ***2.2.11.1 Machine Learning (ML) Research Methodology***

The water quality dataset obtained from the Aik-Stream was initially divided into two subsets: 80% for the training phase, which needed to be sufficiently representative, and 20% for the testing phase to assess the predictive performance of our chosen models. It is important to emphasize that these datasets were randomly distributed. Consequently, all input combinations of water quality parameters were normalized within the range from (0.1 to 0.9). This standardization was carried out to confine the variations in variable values to a predetermined standardized interval. This approach is desirable as it prevents the system from becoming biased towards specific value ranges. Subsequently, we created nineteen distinct



input combinations labeled from 1 to 19 to analyze the influences of input variables on each trained model (GB and RF). This evaluation was conducted using the "Wrapper" approach outlined by (Kohavi & John, 1997) (as detailed in Table 2.3). This method prescribes the selection of input combinations in advance, intending to streamline the calculations by reducing the number of variables considered in the model. This is crucial as the learning algorithm requires a precision metric to evaluate the quality of the developed models.

Precision measurements are typically employed in conjunction with a top-down sequential search algorithm. A variable is eliminated at each process stage, and whether the model maintains adequate precision is verified. This algorithm continues until no more variables are removed or the accuracy falls below an acceptable threshold, as outlined in Table 2.3. The previous researcher used a different parameter combination method in this study (Mehdizadeh, Fathian, Safari, & Khosravi, 2020; Mokhtar et al., 2021).

Table 2. 3 Models developed from the different combinations of the water parameters

Models	Variables																		
1	COD	TOC	OG	NH <sub>3</sub> N	As	Ni	Zn	Cd	Cr	Cl	BOD	TDS	TSS	Pb	pH	T	Hg	NO <sub>3</sub> N	Cu
2	COD	TOC	OG	NH <sub>3</sub> N	As	Ni	Zn	Cd	Cr	Cl	BOD	TDS	TSS	Pb	pH	T	Hg	NO <sub>3</sub> N	
3	COD	TOC	OG	NH <sub>3</sub> N	As	Ni	Zn	Cd	Cr	Cl	BOD	TDS	TSS	Pb	pH	T	Hg		
4	COD	TOC	OG	NH <sub>3</sub> N	As	Ni	Zn	Cd	Cr	Cl	BOD	TDS	TSS	Pb	pH	T			
5	COD	TOC	OG	NH <sub>3</sub> N	As	Ni	Zn	Cd	Cr	Cl	BOD	TDS	TSS	Pb	pH				
6	COD	TOC	OG	NH <sub>3</sub> N	As	Ni	Zn	Cd	Cr	Cl	BOD	TDS	TSS	Pb					
7	COD	TOC	OG	NH <sub>3</sub> N	As	Ni	Zn	Cd	Cr	Cl	BOD	TDS	TSS						
8	COD	TOC	OG	NH <sub>3</sub> N	As	Ni	Zn	Cd	Cr	Cl	BOD	TDS							
9	COD	TOC	OG	NH <sub>3</sub> N	As	Ni	Zn	Cd	Cr	Cl	BOD								
10	COD	TOC	OG	NH <sub>3</sub> N	As	Ni	Zn	Cd	Cr	Cl									
11	COD	TOC	OG	NH <sub>3</sub> N	As	Ni	Zn	Cd	Cr										
12	COD	TOC	OG	NH <sub>3</sub> N	As	Ni	Zn	Cd											
13	COD	TOC	OG	NH <sub>3</sub> N	As	Ni	Zn												
14	COD	TOC	OG	NH <sub>3</sub> N	As	Ni													
15	COD	TOC	OG	NH <sub>3</sub> N	As														
16	COD	TOC	OG	NH <sub>3</sub> N															
17	COD	TOC	OG																
18	COD	TOC																	
19	COD																		

#### 2.2.11.2 Evaluation Metrics

During this study, we compared the observed WQI data and the values predicted by our models. To assess the accuracy of our models, we employed several statistical metrics, namely Mean Absolute Error (MAE) (equations 2.5), Root Relative Squared Error (RRSE) (equations 2.9), Root Mean Square Error (RMSE) (equations 2.10), and the coefficient of



determination ( $R^2$ ) (equations 2.11). These specific metrics were chosen based on the recommendations of previous studies (Malone, Minasny, & McBratney, 2017). The parameters are defined as: "WQI actual" represents the observed or actual WQI value, while "WQI predicted" represents the simulated or predicted WQI value.

$$MAE = \left(\frac{1}{n}\right) * \Sigma |WQI_{pred} - WQI_{actual}| \dots \dots \dots (2.8)$$

$$RRSE = \left(\frac{1}{n}\right) * \Sigma (WQI_i^P - WQI_i^A)^2 \dots \dots \dots (2.9)$$

$$RMSE = \text{sqrt} \left( \left(\frac{1}{n}\right) * \Sigma (WQI_{pred} - WQI_{actual})^2 \right) \dots \dots \dots (2.10)$$

$$R^2 = 1 - \left( \frac{\Sigma (WQI_{actual} - WQI_{pred})^2}{\Sigma (WQI_{actual} - \text{mean}(WQI_{actual}))^2} \right) \dots \dots \dots (2.11)$$

A higher R-squared value indicates a more robust correlation or fit between the observed and actual values. On the other hand, lower values of MAE, MSE, and RMSE indicate improved model performance (Chicco, Warrens, & Jurman, 2021). These evaluation metrics were employed to evaluate how well the regression models predicted the WQI.

### 2.2.11.3 Feature selection

Feature selection is commonly characterized as a search operation to identify a pertinent subset of attributes from the original set. Various approaches to feature selection exist, with our study specifically employing the Recursive Feature Elimination-Linear (RFE-L) method. RFE-L is a widely adopted algorithm for identifying the most relevant features in predictive modeling, as it utilizes a backward selection process to determine the optimal feature combination for predicting the target variable (F. Akhtar et al., 2020; Ebrahimi-Khusfi, Nafarzadegan, & Dargahian, 2021). Initially, the algorithm constructs a model using all available features and computes the importance of each feature within the model. Subsequently, it ranks these features and systematically eliminates the least significant ones based on the model's evaluation metrics, such as RMSE and  $R^2$  (Bagherzadeh, Mehrani, Basirifard, & Roostaei, 2021). The model is then retrained, and the importance of the independent variables is reassessed. This iterative process continues until a specific number of predictive subsets are identified, enabling the assessment or selection of the subset of predictor variables, which are the water quality parameters in this context. The size of the subset is determined to choose the most optimal predictor variables (Kuhn & Johnson, 2018). It is important to note that in this algorithm, the ideal combination of features is achieved



when the values of RMSE approach 0 and  $R^2$  approach 1. This signifies the best model fit for the given dataset.

## 2.3 Results

Descriptive statistics encompassing mean, minimum, maximum values, standard deviation, kurtosis, skewness, and permissible limits for each parameter are detailed comprehensively. The parameters studied include pH, Temperature (T, °C), Electrical Conductivity (EC,  $\mu\text{S}/\text{cm}$ ), Total Dissolved Solids (TDS, mg/L), Total Suspended Solids (TSS, mg/L), Biological Oxygen Demand (BOD, mg/L), Chemical Oxygen Demand (COD, mg/L), Ammonia Nitrogen ( $\text{NH}_3\text{-N}$ , mg/L), Total Organic Carbon (TOC, mg/L), Chloride ( $\text{Cl}^-$ , mg/L), Nitrate Nitrogen ( $\text{NO}_3\text{-N}$ , mg/L), Oil and Grease (O&G, mg/L), and metals such as Zinc (Zn), Nickel (Ni), Copper (Cu), Chromium (Cr), Lead (Pb), Arsenic (As), and Mercury (Hg) shown in Table 2.4. The results showed a slight difference in the pH of two consecutive sampling sites, ranging from 6.15 to 8.95. The temperature stayed constant around 28.00 °C. The various forms of nitrogen content closely mirror the patterns observed in BOD and COD, indicating that elevated nitrate-nitrogen ( $\text{NO}_3\text{-N}$ ) and ammonia-nitrogen ( $\text{NH}_3\text{-N}$ ) concentrations are associated with increased pollutant loads. The total organic carbon content oscillates from 0 to 90 mg/L. As regards the amount of TDS, excessive values from 439 mg/L to 1340 mg/L were observed, and in TSS, from 201 to 334 mg/L. The minimum value of chloride recorded is 74.22. Noticeable differences in the measurements of heavy metals were observed if moves from the upstream region (UP-St) to the midstream (M-St) and downstream (D-St) areas. In this transition, all parameters exhibit a distinct rise at the mid-stream region, except pH, temperature (T), and total suspended solids (TSS), which display no significant differences.



Table 2.4 Descriptive statistics of water quality parameters of the Aik-Stream, including sample size [N = 150 (3x)]

Parameters	Mean			Standard deviation			Kurtosis			Skewness			Permissible Limit
	U-St	M-St	D-St	U-St	M-St	D-St	U-St	M-St	D-St	U-St	M-St	D-St	HO (mg/L)
pH	7.49	6.18	7.94	0.78	0.767	1.295	-0.09	-1.21	-0.02	-0.32	-0.31	-1.09	6.5-7
COD mg/L	6.054	62.39	43.15	6.05	16.972	15.62	-0.62	0.65	1.30	1.52	0.02	1.46	250
BOD mg/L	12.02	43.02	24.56	12.0	6.733	9.55	-0.34	-0.05	0.36	-0.48	-0.46	-0.55	50
TDS mg/L	6.24	7.00	6.87	0.42	0.000	0.33	1.28	0.00	-2.34	0.36	0.00	3.63	300-600
TSS mg/L	5.94	6.00	0.24	0.00	0.000	0.46	-0.72	0.00	0.07	-0.47	0.00	-0.39	250
NH <sub>3</sub> -N mg/L	5.96	22.73	4.424	2.56	3.447	4.42	0.07	-0.64	-0.83	-1.53	-0.50	-1.36	50
O&G mg/L	2.12	5.75	2.19	0.81	1.309	1.06	-0.22	-0.50	0.60	-1.46	-0.24	-0.36	10
TOC mg/L	13.6	110.9	40.33	4.55	17.512	28.12	-0.54	-0.28	0.65	0.91	-0.96	-1.31	15
NO <sub>3</sub> -N mg/L	0.45	0.51	0.17	0.50	0.505	0.37	0.20	-0.04	1.84	-2.04	-2.08	1.47	15
Cl <sup>-</sup> mg/L	157.4	221.2	151.6	44.2	28.915	45.15	-0.12	-0.30	0.46	-0.72	-0.48	-0.69	200
Cu mg/L	0.47	0.50	0.58	0.50	0.140	0.49	.121	-7.14	-0.34	-2.06	51.0	-1.96	1
Zn mg/L	-0.14	1.00	0.23	0.34	0.316	0.42	-2.173	-0.46	1.33	2.83	8.02	-0.24	3
Cr mg/L	0.20	1.94	1.27	0.40	0.947	0.64	1.578	0.70	2.68	0.50	-0.41	7.39	0.05
Pb mg/L	0.76	2.12	1.44	0.42	0.739	0.58	-1.286	0.11	0.93	-0.36	-0.41	-0.06	0.05
Cd mg/L	0.00	0.98	0.40	0.00	0.510	.494	0.00	-0.03	0.44	0.00	1.14	-1.88	0.05
Ni mg/L	0.00	0.90	0.54	0.00	0.300	0.50	0.00	-2.78	-0.17	0.00	5.99	-2.05	0.1
Ar mg/L	0.00	1.53	0.83	0.00	0.504	0.72	0.00	-0.12	0.26	0.00	-2.06	-1.01	0.05
Hg mg/L	0.00	0.20	0.06	0.00	0.401	0.24	0.00	1.57	3.73	0.00	0.50	12.44	0.02

Table 2.5 illustrates the Pearson Correlation matrix, indicating a strong association between various physicochemical parameters and heavy metals. The relationship between COD and BOB (0.69), COD and TDS (0.84), COD and pH (-0.57), COD and NH<sub>3</sub>-N (0.82), COD versus TSS (0.65), COD and O & G (0.82), COD and TOC (0.76), COD and Cl<sup>-</sup> (0.63), COD and Cu (0.67), COD and Zn (0.68), COD and Cr (0.60), COD and Pb (0.61), COD and Cd (0.59), COD and Ni (0.61), COD and As (0.74) and COD versus Hg (0.42). Moreover, BOD and other water parameters had a strong relationship, such as BOD and TDS (0.54), BOD and NH<sub>3</sub>-N (0.83), BOD and Cu (0.71), BOD and As (0.84), DOB and Cr (0.81), and BOD & Zn (0.79). The correlation between the heavy metal concentrations also had a strong linear relationship found in Cu and Zn (0.61), Cu and Cr (0.74), Cu and Pd (0.67), Cr and Ni (0.83), Zn and Cr (0.86). The result indicated a positive and strong relationship between the heavy metal and physiochemical parameters of the water quality.



Table 2.5 Pearson correlation matrix of all studied variables

Variables	COD	BOD	TDS	pH	T	NH <sub>3</sub> -N	TSS	OG	TOC	NO <sub>3</sub> -N	Cl	Cu	Zn	Cr	Pb	Cd	Ni	As	Hg
<b>COD</b>	1																		
<b>BOD</b>	0.69**	1																	
<b>TDS</b>	0.84**	0.57*	1																
<b>pH</b>	-0.57*	-0.56*	-0.6*	1															
<b>T</b>	0.19	0.42**	-0.03	-0.1	1														
<b>NH<sub>3</sub>-N</b>	0.82**	0.83**	0.74**	-0.64*	0.43*	1													
<b>TSS</b>	0.65**	0.63*	0.59*	-0.52*	0.24	0.71**	1												
<b>OG</b>	0.82**	0.75**	0.80**	-0.69*	0.17	0.83**	0.79*	1											
<b>TOC</b>	0.76**	0.82**	0.72**	-0.73*	0.28	0.87**	0.86**	0.87**	1										
<b>NO<sub>3</sub>-N</b>	0.02	-0.12	0.13	0.12	-0.16	-0.12	0.09	0.07	-0.04	1									
<b>Cl</b>	0.63**	0.69**	0.55**	-0.60*	0.38*	0.76**	0.83**	0.82**	0.85**	0.16	1								
<b>Cu</b>	0.67**	0.71*	0.61**	-0.55*	0.31*	0.78**	0.66*	0.72*	0.76*	-0.15	0.61*	1							
<b>Zn</b>	0.68**	0.79*	0.68**	-0.58*	0.21	0.76**	0.55*	0.70*	0.78**	0.03	0.67*	0.61*	1						
<b>Cr</b>	0.60**	0.81**	0.53*	-0.57*	0.49*	0.84**	0.53**	0.69*	0.76**	-0.07	0.72*	0.66*	0.71*	1					
<b>Pb</b>	0.61**	0.68*	0.58*	-0.49*	0.33*	0.76**	0.49*	0.78**	0.66*	0.02	0.60*	0.69*	0.62**	0.74*	1				
<b>Cd</b>	0.59**	0.76**	0.55*	-0.54*	0.48*	0.82**	0.48*	0.71*	0.69*	-0.02	0.68*	0.59*	0.70*	0.85*	0.76*	1			
<b>Ni</b>	0.61**	0.73*	0.51**	-0.59*	0.58**	0.79**	0.69*	0.71*	0.79**	-0.05	0.71*	0.66*	0.66*	0.71*	0.69*	0.69*	1		
<b>As</b>	0.74**	0.84**	0.73*	-0.70**	0.28	0.88**	0.71**	0.86**	0.90*	0.02	0.82**	0.73**	0.86**	0.83**	0.74*	0.80**	0.72**	1	
<b>Hg</b>	0.42**	0.42*	0.42*	-0.45*	0.27	0.55*	0.57*	0.57*	0.56*	-0.07	0.50*	0.61*	0.36*	0.42*	0.54*	0.45*	0.62*	0.52*	1

\*\* Correlation coefficient significant p value at 0.01; \* Correlation coefficient p value significant at p value 0.05



### ***2.3.1 Classification of the sampling sites***

The Hierarchical Agglomerative Cluster Analysis (HACA) was applied to the Aik-Stream wastewater quality dataset to assess spatial differences across sampling sites. This analysis led to classifying sites into three distinct groups, as depicted in Fig. 2.7. Each group consists of sites with analogous water quality parameters. These groups were labeled based on the extent of pollution they exhibited, including the Highly Polluted Zone (HPZ), Moderately Polluted Zone (MPZ), and Less Polluted Zone (LPZ). Sites from 1-51 were allocated to the LPZ category. Situated in the upper reaches of the Aik-Stream, these locations exhibited minimal levels of contaminants, including COD, BOD, NH<sub>3</sub>-N, EC, TDS, NO<sub>3</sub>-N, Cl<sup>-</sup>, Pb, Cr, Ni, As, Hg, Cu, Cd, and Zn. As they are upstream from Sialkot city, these areas experience limited impact from industrial discharges and urban sewage. Sites from 151-117 were categorized under the HPZ and situated near the mid-stream of Aik-Stream. This zone, considered the most polluted, showed elevated COD, BOD, EC, NH<sub>3</sub>-N, TDS, NO<sub>3</sub>-N, Cl<sup>-</sup>, Pb, Cr, Ni, As, Hg, Cu, Cd, and Zn levels. The increased pollution levels here can be attributed to effluents from industries in the mid-stream region. Lastly, sites 118-150 were grouped under the MPZ in the downstream area. This region has fewer industries than the mid-stream, showcasing moderate pollution levels.

Cluster analysis was also used to show the similarity among the parameters that cause water pollution. Cluster one includes the parameters whose value is higher than the permissible level set by WHO (2004). Cluster 1 contained the following variables: Cd, As, COB, Cu, NH<sub>3</sub>-N, Pb, BOD and TOC. Cluster 2 includes parameters such as TDS, TSS, Zn, Ni, and Cl<sup>-</sup> the concentration of this variable is lower than in the high-polluted zone and higher than in the less-polluted zones. Cluster three, the less polluted zone, contains the following parameters: pH, Cr, and O&G (Fig 2.8).



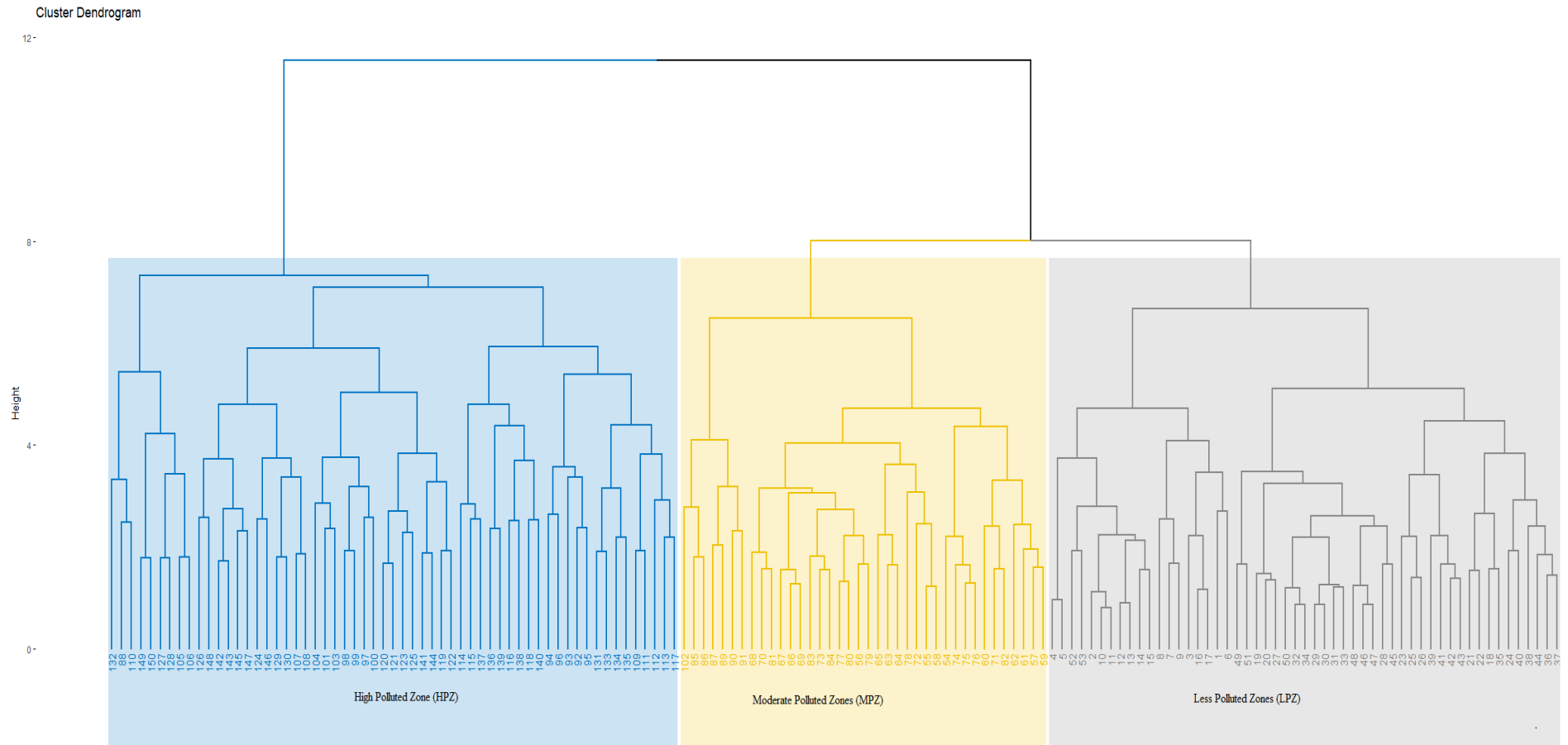


Figure 2. 7 Dendrogram showing different clusters of sampling zones located in Aik Stream



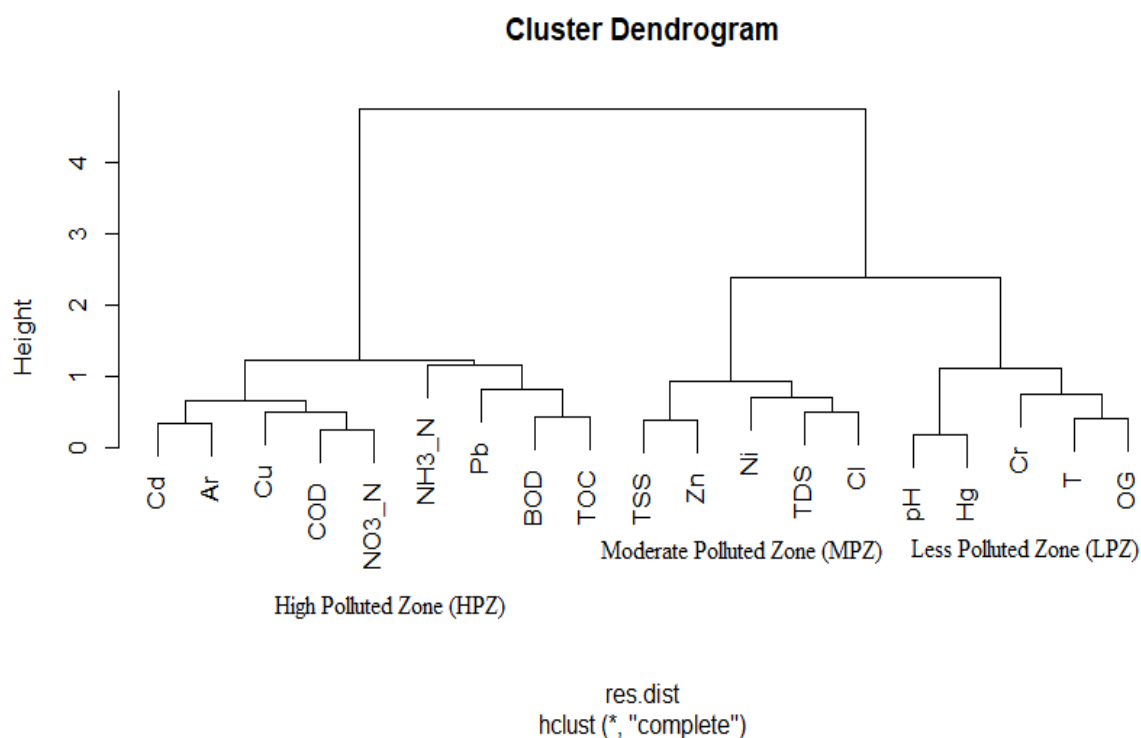


Figure 2. 8 Dendrogram showing the concentration of parameters in different sites located in Aik Stream.

The concentration of heavy metal and physiochemical parameters in the Aik stream and its comparison with the three zones are depicted in Table 2.6 and Fig 2.9 & 2.10. The result indicated that the concentration of heavy metal is in the following order:  $Cu > Zn > Cr > Pb > Cd > Ni > As > Hg$  in (mg/l). And the physiochemical parameters follow this order:  $COD > TOC > BOD > TDS > TSS > pH > Cl^- > NH_3-N > NO_3-N > O\&G$  mg/l. The concentration of Cu ranged between (1.96 to 2.56), the highest concentration of Cu, 2.56, was recorded in the highly polluted zone, and the lowest concentration of Cu was observed in the less polluted zone, 1.96 mg/l. Zn concentration in water varied from (0.95 – 2.74), and the highest concentration of Zn was recorded (2.74) in a highly polluted zone. The minimum concentration of Zn (0.95 mg/l) was noted at a less polluted zone in the stream. The average concentration of Cr (0.98 mg/l) was observed in all three different zones. A higher concentration was noted (0.86 to 2.01 mg/l) in a highly polluted zone during monsoon. The Cd ranged from (0.05 to 0.90) in the Aik stream's less and high polluted zone. The concentration of Ni ranged from (0.03 to 0.83), the highest concentration of Ni was noted in the higher pollutant zone (0.83), and a lesser concentration was recorded in LPZ (0.03 mg/l). The average concentration of As and Hg varied between (0.34 and 0.21) in all 150 water samples collected from different sites of the Aik Stream.



Table 2.6 Zone-wise descriptive statistics among all parameters of the Aik Stream.

	ZONE	N	MEAN	SD	SE
<b>COD</b>	HPZ	66	62.1212	16.3533	2.01296
	LPZ	51	19.0980	6.0539	0.84772
	MPZ	33	34.9394	6.5999	1.14890
<b>BOD</b>	HPZ	66	41.3182	7.2008	0.88635
	LPZ	51	25.9804	12.0208	1.68325
	MPZ	33	19.5758	6.1493	1.07046
<b>TDS</b>	HPZ	66	1019.5000	219.3420	26.99914
	LPZ	51	609.5294	72.1849	10.10790
	MPZ	33	779.0303	125.5783	21.86037
<b>PH</b>	HPZ	66	6.3591	0.6128	0.07543
	LPZ	51	7.4855	0.7595	0.10635
	MPZ	33	8.4645	0.8319	0.14482
<b>T</b>	HPZ	66	28.7545	0.0502	0.00618
	LPZ	51	28.7000	0.0000	0.00000
	MPZ	33	28.7727	0.0452	0.00787
<b>NH<sub>3</sub>-N</b>	HPZ	66	22.2424	3.3331	0.41028
	LPZ	51	5.9608	2.5687	0.35970
	MPZ	33	13.1515	2.9168	0.50775
<b>TSS</b>	HPZ	66	311.5000	26.9953	3.32289
	LPZ	51	261.6078	31.7692	4.44858
	MPZ	33	245.2727	18.4412	3.21021
<b>O &amp; G</b>	HPZ	66	5.1867	1.4771	0.18182
	LPZ	51	2.1471	0.6970	0.09759
	MPZ	33	1.6252	0.5488	0.09553
<b>TOC</b>	HPZ	66	103.8485	20.4264	2.51432
	LPZ	51	13.6863	4.5541	0.63770
	MPZ	33	22.4242	9.3576	1.62895
<b>NO<sub>3</sub>-N</b>	HPZ	66	0.4506	0.2848	0.03505
	LPZ	51	0.5194	0.3217	0.04505
	MPZ	33	0.3545	0.1563	0.02721
<b>Cl<sup>-</sup></b>	HPZ	66	218.9206	27.4045	3.37326
	LPZ	51	157.5661	44.2808	6.20056
	MPZ	33	125.0988	21.8203	3.79843
<b>Cu</b>	HPZ	66	2.5668	0.6899	0.08492
	LPZ	51	1.9632	0.8847	0.12388
	MPZ	33	2.0238	0.7800	0.13578
<b>Zn</b>	HPZ	66	2.7407	0.7525	0.09263
	LPZ	51	0.9527	0.2327	0.03258
	MPZ	33	1.2083	0.1765	0.03073
<b>Cr</b>	HPZ	66	1.9811	0.8311	0.10230
	LPZ	51	0.2999	0.1918	0.02686
	MPZ	33	1.1722	0.1509	0.02627
<b>Pb</b>	HPZ	66	2.0154	0.7197	0.08859
	LPZ	51	0.8634	0.3225	0.04517
	MPZ	33	1.3937	0.3049	0.05308
<b>Cd</b>	HPZ	66	0.9073	0.4024	0.04953
	LPZ	51	0.0510	0.0166	0.00233
	MPZ	33	0.3703	0.2051	0.03570
<b>Ni</b>	HPZ	66	0.8392	0.3128	0.03851
	LPZ	51	0.0351	0.0102	0.00143
	MPZ	33	0.5247	0.1817	0.03162
<b>As</b>	HPZ	66	1.4868	0.3713	0.04570
	LPZ	51	0.1591	0.0591	0.00827
	MPZ	33	0.4951	0.2123	0.03696
<b>Hg</b>	HPZ	66	0.3429	0.2469	0.03039
	LPZ	51	0.0271	0.0129	0.00181
	MPZ	33	0.2935	0.1122	0.01954



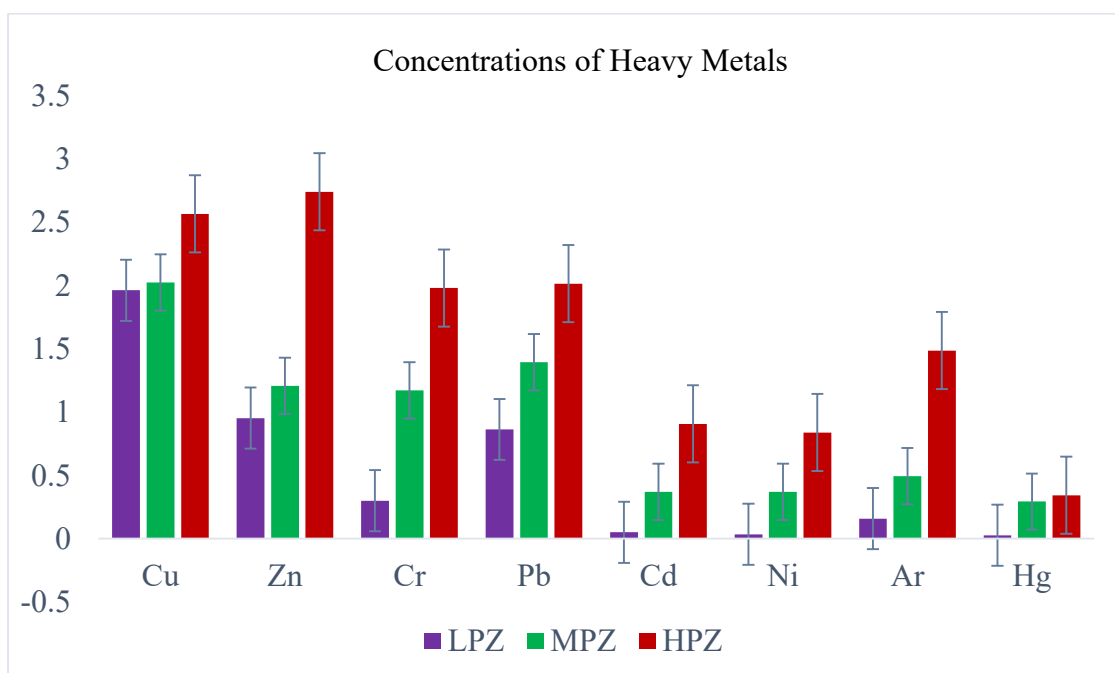


Figure 2. 9 Graph showing the concentration of heavy metal in three zones (LPZ, MPZ, and HPZ)

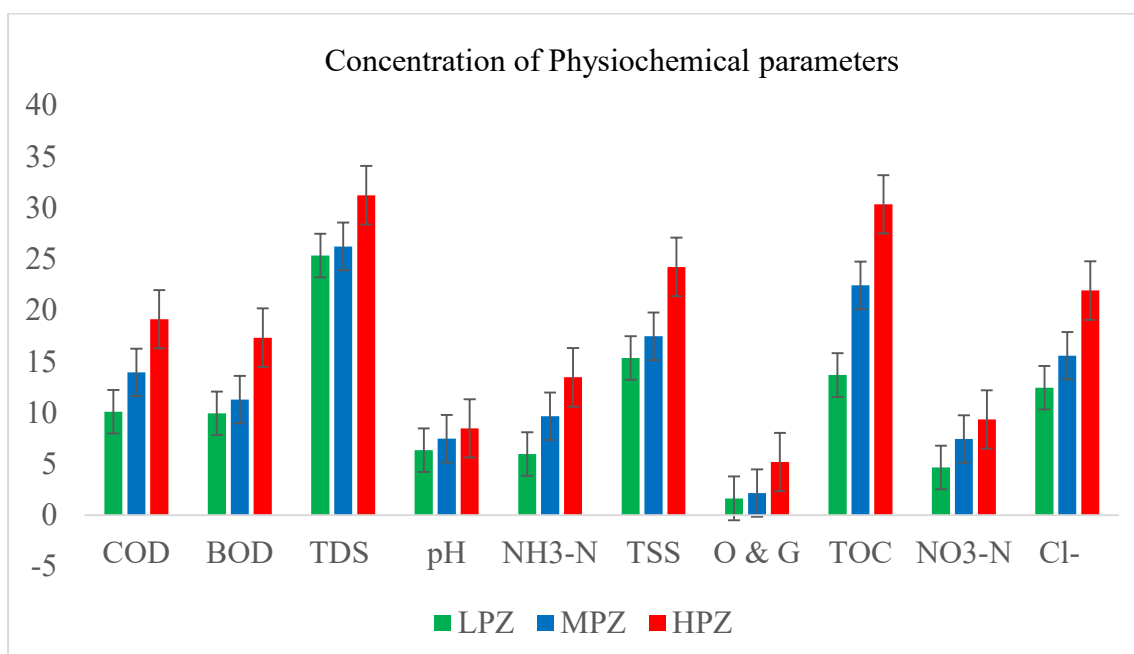


Figure 2. 10 Graph showing the physicochemical characteristics of Aik stream water quality in three zones

### 2.3.2 Variation among the zones and within the parameters

A one-way analysis of variance (ANOVA) was used to confirm our results further and indicate the mean variation among the groups and within the parameters. It further clarifies



whether there are statistically meaningful variations among the groups. A p-value less than a predefined significance level (often denoted as  $\alpha$ ) typically indicates significant differences between at least two groups. In (Table 2.7 and Fig 2.11 (A-S)) water quality variables, such as COD, BOD, TDS, pH, and so on, have been subjected to One-Way ANOVA. The p-values for each parameter are consistently reported as "0.001," which suggests highly significant differences among the groups for each parameter. The results indicate that the different groups or categories under investigation exhibit statistically significant parameter value variations. The box plot shown in Fig. 2.11, spanning panels A to S, provides a graphical depiction of the variability of variables across three distinct zones. Researchers can use these findings to make informed conclusions about the impact of the group or category on the parameter in question. The ANOVA test is handy when dealing with datasets that violate the assumption of equal variances, providing robust and reliable results for data analysis.

Statistical analysis is the most fundamental step to understanding the nature of the data. So, for this purpose, various Normality Tests are applied to understand whether data is standard. Table 2.8 contains a detailed overview of different normality tests. The table shows the results of three normality tests for each parameter: Shapiro-Wilk, Kolmogorov-Smirnov, and Anderson-Darling. We used the Shapiro-Wilk test to investigate whether the dataset follows a normal distribution. It is susceptible to checking the normality of the parameter. If the p-value is less than 0.05, the test suggests that the data is normally distributed. Shapiro-Wilk test results indicate that parameters such as (pH, TSS, and As) have p-value less than 0.05, indicating that they are not normally distributed while the rest are normally distributed. Kolmogorov-Smirnov test is another normality test used to evaluate the normality of the variables. It relates the observed distribution of the data to the hypothetical normal distribution. The outcome contains the test statistics and p-value. Like the Shapiro-Wilk test, a lesser p-value shows non-normality. For example, "Ni," "Zn," and "Cd" have lesser p-values, indicating deviations from normality. We used the third normality test, the Anderson-Darling Test, to check the distribution of the parameters. Table 3.6 shows the result of the Shapiro-Wilk, Kolmogorov-Smirnov, and Anderson-Darling Test containing t-static and p-value. These tests also follow the above two tests if the p-value is significant, it indicates that the parameters are not normal. Parameters such as "Cd," "Ni," and "Hg" exhibit low p-values, suggesting non-normal distributions. Overall, these normality tests provide valuable information about the data distribution for each parameter. Non-normal data can affect the choice of statistical methods used for analysis. Researchers often consider transformations or



non-parametric tests when dealing with non-normally distributed data to ensure the validity of their statistical inferences.

Table 2.7 Mean deviation of the water quality in different sites of the Aik Stream.

<b>Zones</b>	<b>Variables</b>	<b>Mean difference<math>\pm</math> Std. Error</b>
<b>LPZ</b>	COD	19.0980 $\pm$ 0.84772
	BOD	25.9804 $\pm$ 1.68325
	TDS	609.5294 $\pm$ 10.10790
	pH	7.4855 $\pm$ 0.10635
	T	28.7000 $\pm$ 0.00000
	NH <sub>3</sub> -N	5.9608 $\pm$ 0.35970
	TSS	261.6078 $\pm$ 4.44858
	O & G	2.1471 $\pm$ 0.09759
	TOC	13.6863 $\pm$ 0.63770
	NO <sub>3</sub> -N	0.5194 $\pm$ 0.04505
	Cl <sup>-</sup>	157.5661 $\pm$ 6.20056
	Cu	1.9632 $\pm$ 0.12388
	Zn	0.9527 $\pm$ 0.9527
	Cr	0.2999 $\pm$ 0.2999
	Pb	0.8634 $\pm$ 0.04517
	Cd	0.0510 $\pm$ 0.00233
	Ni	0.0351 $\pm$ 0.00143
	As	0.1591 $\pm$ 0.00827
	Hg	0.0271 $\pm$ 0.00181
<b>MPZ</b>	COD	34.9394 $\pm$ 1.14890
	BOD	19.5758 $\pm$ 1.07046
	TDS	779.0303 $\pm$ 21.86037
	pH	8.4645 $\pm$ 0.14482
	T	28.7727 $\pm$ 0.00787
	NH <sub>3</sub> -N	13.1515 $\pm$ 0.50775
	TSS	245.2727 $\pm$ 3.21021
	O & G	1.6252 $\pm$ 0.09553
	TOC	22.4242 $\pm$ 1.62895
	NO <sub>3</sub> -N	0.3545 $\pm$ 0.02721
	Cl <sup>-</sup>	125.0988 $\pm$ 3.79843
	Cu	2.0238 $\pm$ 0.13578
	Zn	1.2083 $\pm$ 0.03073
	Cr	1.1722 $\pm$ 0.02627
	Pb	1.3937 $\pm$ 0.05308
	Cd	0.3703 $\pm$ 0.03570
	Ni	0.5247 $\pm$ 0.03162
	As	0.4951 $\pm$ 0.03696
	Hg	0.2935 $\pm$ 0.01954
<b>HPZ</b>	COD	62.1212 $\pm$ 2.01296
	BOD	41.3182 $\pm$ 0.88635
	TDS	1019.5 $\pm$ 26.999
	pH	6.3591 $\pm$ 0.07543
	T	28.7545 $\pm$ 0.00618
	NH <sub>3</sub> -N	22.2424 $\pm$ 0.41028
	TSS	311.500 $\pm$ 3.32289
	O & G	5.1867 $\pm$ 0.18182
	TOC	103.8485 $\pm$ 2.51432
	NO <sub>3</sub> -N	0.4506 $\pm$ 0.03505
	Cl <sup>-</sup>	218.9206 $\pm$ 3.37326



	Cu	2.5668 ± 0.08492
	Zn	2.7407 ± 0.09263
	Cr	1.9811 ± 0.10230
	Pb	2.0154 ± 0.08859
	Cd	0.9073 ± 0.04953
	Ni	0.8392 ± 0.03851
	As	1.4868 ± 0.04570
	Hg	0.3429 ± 0.03039

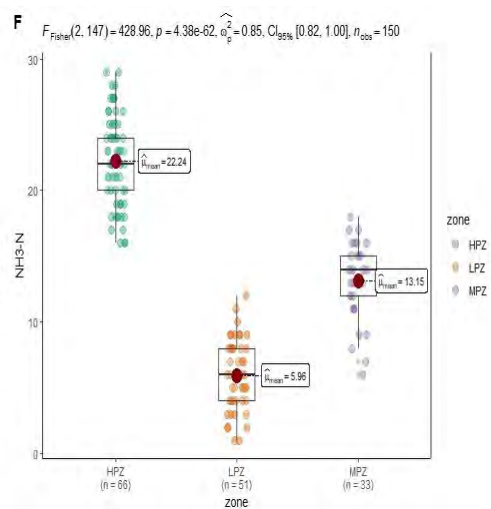
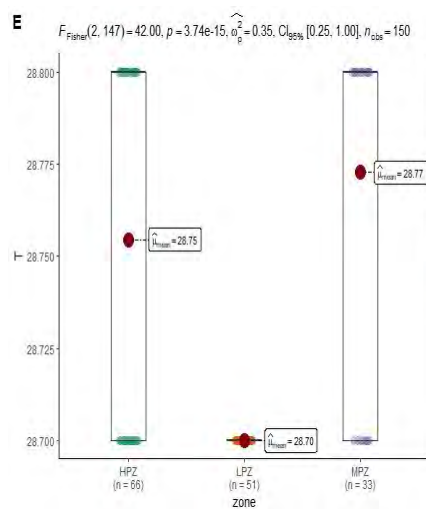
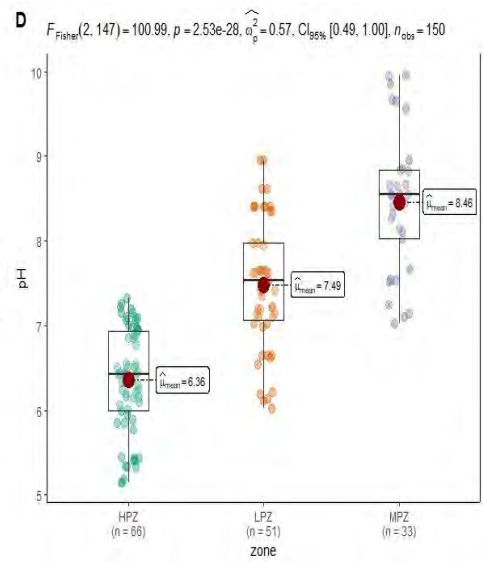
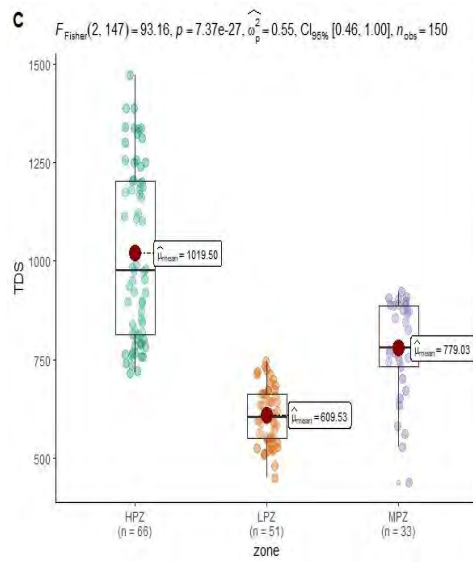
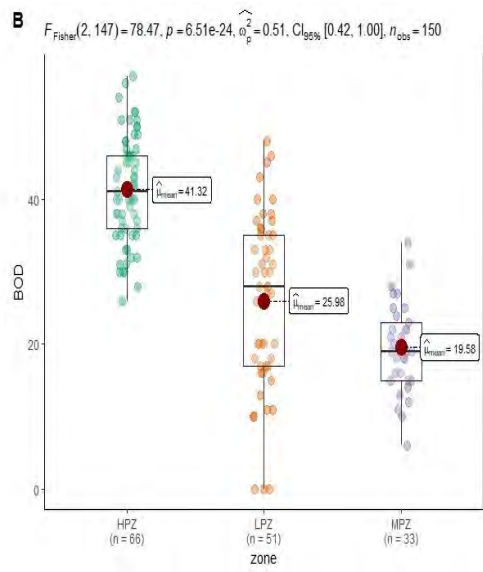
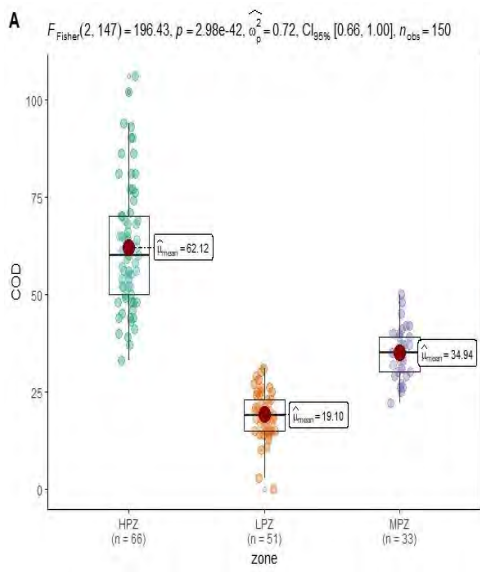
Table 2.8 Normality Tests for all the studied water quality parameters.

	Tests	statistic	p-value
<b>COD</b>	Shapiro-Wilk	0.955	< .001
	Kolmogorov-Smirnov	0.0994	0.103
	Anderson-Darling	1.855	< .001
<b>BOD</b>	Shapiro-Wilk	0.990	0.331
	Kolmogorov-Smirnov	0.0494	0.858
	Anderson-Darling	0.242	0.767
<b>TDS</b>	Shapiro-Wilk	0.987	0.198
	Kolmogorov-Smirnov	0.0500	0.847
	Anderson-Darling	0.405	0.348
<b>pH</b>	Shapiro-Wilk	0.980	0.026
	Kolmogorov-Smirnov	0.0725	0.409
	Anderson-Darling	0.806	0.036
<b>T</b>	Shapiro-Wilk	0.851	< .001
	Kolmogorov-Smirnov	0.2400	< .001
	Anderson-Darling	8.670	< .001
<b>NH<sub>3</sub>-N</b>	Shapiro-Wilk	0.989	0.316
	Kolmogorov-Smirnov	0.0668	0.516
	Anderson-Darling	0.455	0.265
<b>TSS</b>	Shapiro-Wilk	0.994	0.785
	Kolmogorov-Smirnov	0.0567	0.721
	Anderson-Darling	0.262	0.699
<b>O &amp; G</b>	Shapiro-Wilk	0.987	0.184
	Kolmogorov-Smirnov	0.0671	0.508
	Anderson-Darling	0.650	0.088
<b>TOC</b>	Shapiro-Wilk	0.965	< .071
	Kolmogorov-Smirnov	0.1230	0.021
	Anderson-Darling	2.214	< .031
<b>NO<sub>3</sub>-N</b>	Shapiro-Wilk	0.975	0.028
	Kolmogorov-Smirnov	0.0614	0.623
	Anderson-Darling	0.819	0.034
<b>Cl</b>	Shapiro-Wilk	0.989	0.275
	Kolmogorov-Smirnov	0.0547	0.761
	Anderson-Darling	0.407	0.346
<b>Cu</b>	Shapiro-Wilk	0.957	< .001
	Kolmogorov-Smirnov	0.1208	0.025
	Anderson-Darling	2.173	< .001
<b>Zn</b>	Shapiro-Wilk	0.921	< .001
	Kolmogorov-Smirnov	0.1248	0.039
	Anderson-Darling	3.924	< .111
<b>Cr</b>	Shapiro-Wilk	0.918	< .061
	Kolmogorov-Smirnov	0.1257	0.017



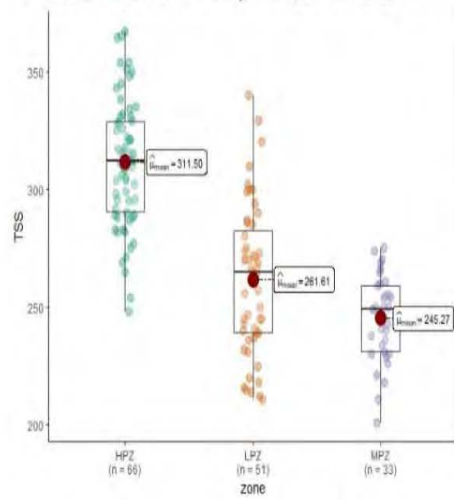
<b>Pb</b>	Anderson-Darling	3.935	< .021
	Shapiro-Wilk	0.964	< .031
	Kolmogorov-Smirnov	0.0668	0.515
<b>Cd</b>	Anderson-Darling	1.337	0.082
	Shapiro-Wilk	0.955	< .061
	Kolmogorov-Smirnov	0.1612	< .001
<b>Ni</b>	Anderson-Darling	3.515	< .031
	Shapiro-Wilk	0.928	< .001
	Kolmogorov-Smirnov	0.1726	< .001
<b>As</b>	Anderson-Darling	5.432	< .001
	Shapiro-Wilk	0.975	0.007
	Kolmogorov-Smirnov	0.0829	0.254
<b>Hg</b>	Anderson-Darling	0.966	0.015
	Shapiro-Wilk	0.844	< .031
	Kolmogorov-Smirnov	0.1851	< .041
	Anderson-Darling	6.772	< .051



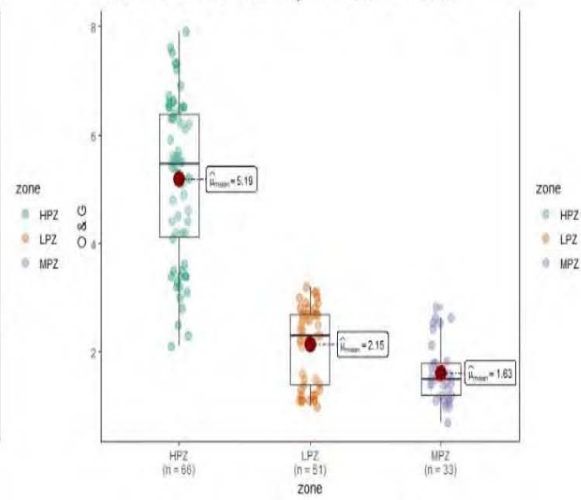




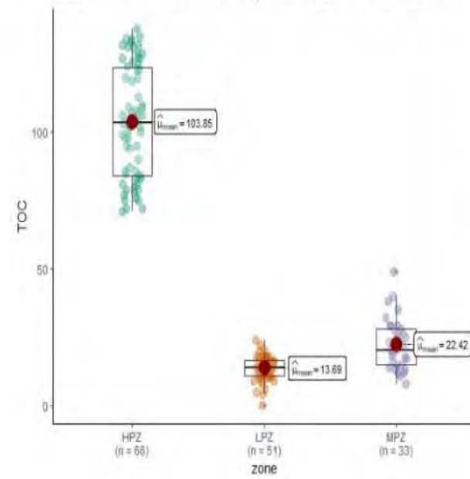
**G**  $F_{\text{Fisher}}(2, 147) = 82.85, p = 8.07e-25, \hat{\sigma}_p^2 = 0.52, CI_{95\%} [0.43, 1.00], n_{\text{obs}} = 150$



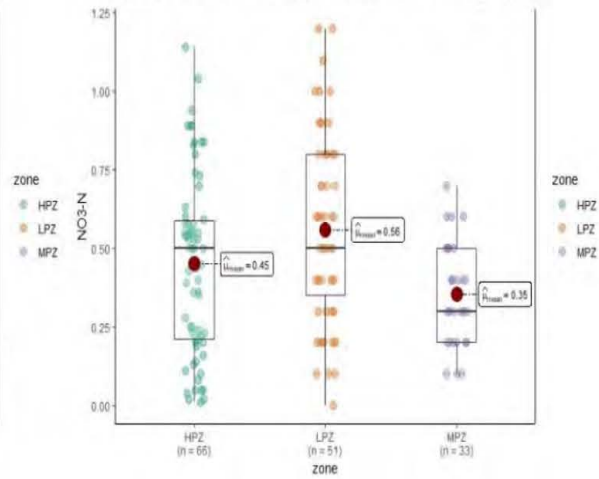
**H**  $F_{\text{Fisher}}(2, 147) = 165.02, p = 2.66e-38, \hat{\sigma}_p^2 = 0.69, CI_{95\%} [0.62, 1.00], n_{\text{obs}} = 150$



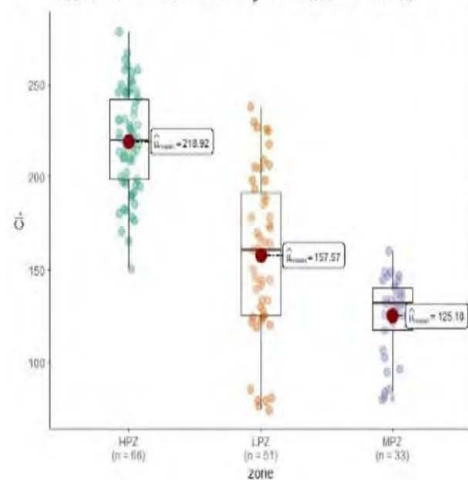
**I**  $F_{\text{Fisher}}(2, 147) = 663.65, p = 2.55e-74, \hat{\sigma}_p^2 = 0.90, CI_{95\%} [0.88, 1.00], n_{\text{obs}} = 150$



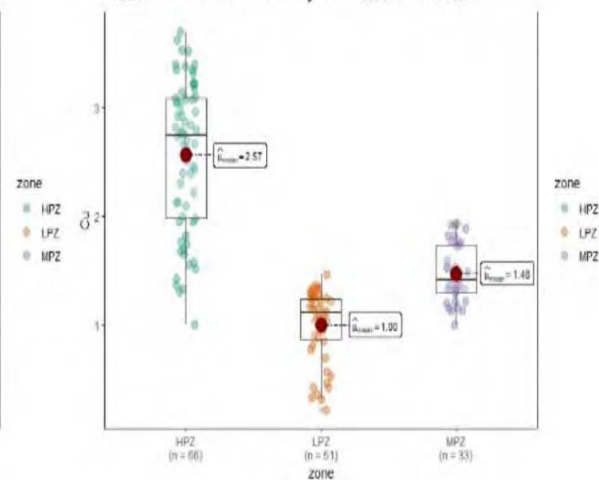
**J**  $F_{\text{Fisher}}(2, 147) = 6.03, p = 3.03e-03, \hat{\sigma}_p^2 = 0.06, CI_{95\%} [9.72e-03, 1.00], n_{\text{obs}} = 150$



**K**  $F_{\text{Fisher}}(2, 147) = 101.62, p = 1.93e-28, \hat{\sigma}_p^2 = 0.57, CI_{95\%} [0.49, 1.00], n_{\text{obs}} = 150$



**M**  $F_{\text{Fisher}}(2, 147) = 142.33, p = 4.12e-36, \hat{\sigma}_p^2 = 0.65, CI_{95\%} [0.58, 1.00], n_{\text{obs}} = 150$





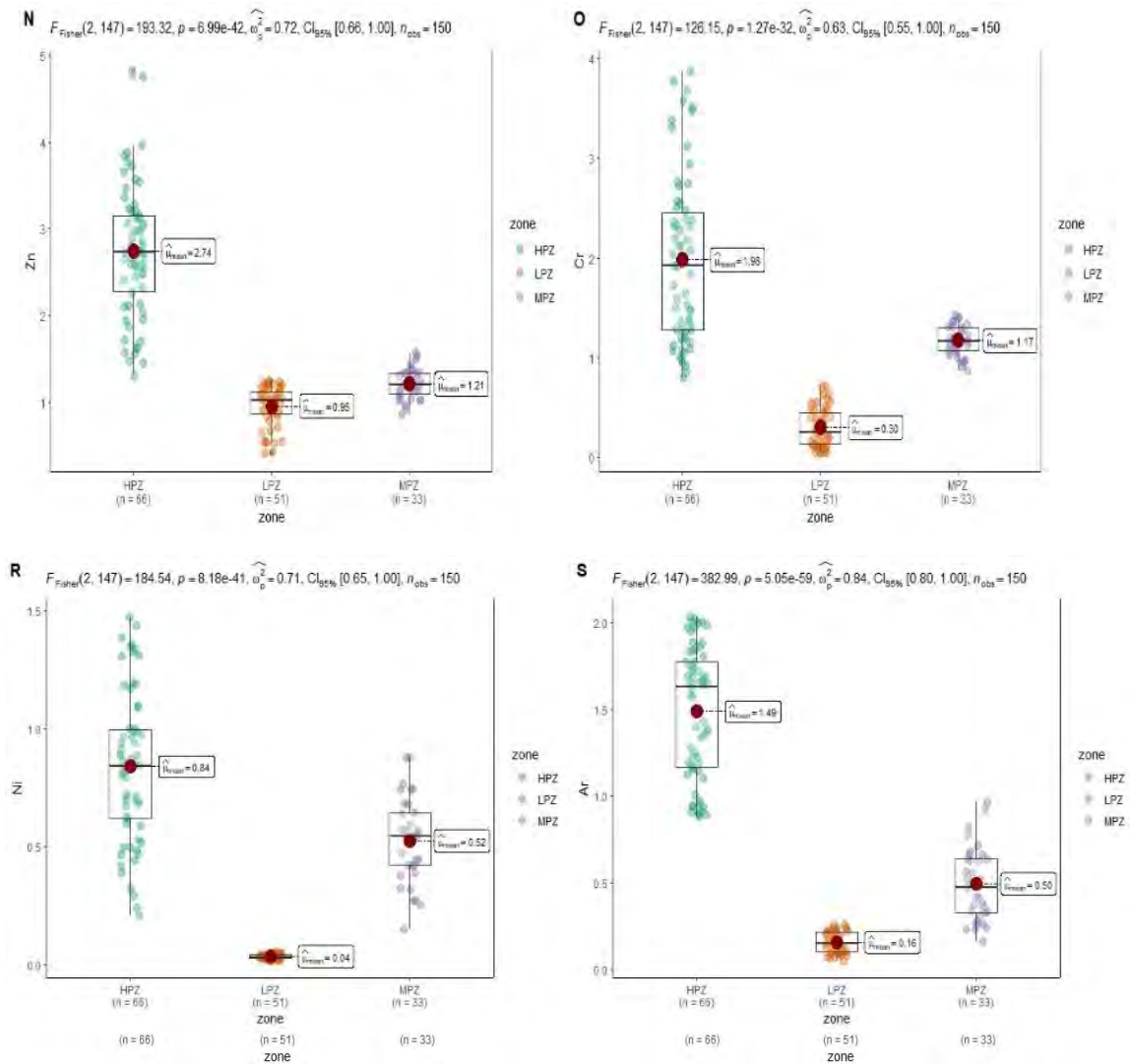


Figure 2. 11 (A-S) Box plot variance comparison of all studied parameters in different pollution gradient zones of Aik-Stream (HPZ, LPZ, and MPZ)

### 2.3.3 Principal Component Analysis (PCA)

PCA was performed based on HCCA in three different zones, HPZ, MPZ, and LPZ, to determine the possible water sources for pollution and specify dominant variables mainly responsible for changing the water quality (Fig. 2.12 and Table 2.9). The six components were primarily responsible for the Aik Stream water quality, the contribution of this component as such: 94.50, 81.2, 22.9, 31.6, 56.7, and 77.3% of the total variation of the water pollution in LPZ, MPZ, and HPZ. The six parameters contain more than one eigenvalue of each parameter; the total variation of the six components is 75% (Table 2.9).

In the case of HPZ, out of six components, component one showed 30.30% to the total variance showed a strong association with COD (0.72), BOD (0.53), T (0.77),  $\text{NH}_3\text{-N}$  (0.82), TSS (0.61), O & G (0.80), TOC (0.77),  $\text{Cl}^-$  (0.53), Pb (0.70), and As (0.84) these



elements mostly comes from the industrial activities, such as tanneries, discharge a large amount of salt such as NaCl, chromium sulfate and CaCO<sub>3</sub> in the stream water. Components two and three described 17% and 8% of the variance. Component two showed a strong and positive relationship on BOD (0.53), O & G (0.30), Cu (0.83), Zn (0.87), Cd (0.46) and Ni (0.41), respectively, and component three represents positive correlation with COD (0.28), Cr (0.66) and Cd (0.50). Component four describes high variation in BOD and Cd, while the fifth component indicates a strong correlation in only Cu (0.38). The component explained 5.45% of the variance to the total variation and showed a strong correlation between NO<sub>3</sub>-N (0.35) and Hg (0.83). Heavy metal and other water-polluted parameters are high in the midstream region.

In the case of MPZ, component one described 30% of the total variance and positively correlated with T (0.60), NH<sub>3</sub>-H (0.82), NO<sub>3</sub>-N (0.42), and TOC (0.77). Component two evaluated 17% of the total variance and indicated the positive association with pH (0.52), O & G (0.30), Cu (0.83), Zn (0.87), and Cd (0.46); these parameters play a prominent role in water quality pollution in Aik Stream. Component three showed a positive relationship with Cr and Cd in MPZ. Component four explained 7% of the total variance and positively correlated with BOD (0.43), TDS (0.20) and Cd (0.55). Component five described 6% of the total variance and indicated a positive and significant correlation with pH (0.77) and Ni (0.48), respectively. Component six shared 5% of the total variance and positively correlated with COD and NO<sub>3</sub>-N.

In LPZ, the heavy metal concentration and other water quality parameters are not too high compared to HPZ and MPZ in the Aik Stream water quality. The first three components shared a total of 17, 10 and 6% of the total variance, and the relationship between the parameters indicates that component one showed positive and significant correlation with pH (0.62), NO<sub>3</sub>-N (0.42), Cu (0.49) and Cd (0.66). In contrast, component two indicates a positive correlation with pH (0.52), TSS (0.32), Cu (0.63), and Cd (0.56), respectively, while the correlation coefficient of component three indicates a positive correlation with TOC and Cd. The last three components shared 7, 6, and 5% of the total variance. Components four and five showed a positive correlation with NO<sub>3</sub>-N (0.22) and Cd (0.55), TDS (0.55), and Ni (0.48).

Table 2.10 gives an overview of the Eigenvalues that come from a Principal Component Analysis (PCA), a method for reducing the dimensionality of the dataset while



retaining the most essential data. The evaluation of each principal component's significance and contribution to explaining the variance in the data is made possible using eigenvalues, which are essential in PCA. The "Component" column in Table 2.10 lists the principal components produced by the PCA, ranked from the most to the least significant. Each eigenvalue's magnitude is shown in the "Eigenvalue" column. The eigenvalues represent the variance that each primary component contributes.

The "% of Variance" column contains information about the proportion of variance that each component explains. The eigenvalues appear to gradually decline as we proceed down the table, which is usual for PCA. The preliminary components naturally acquire the highest variance, while the later components add fewer and fewer. The "Cumulative %" column shows the cumulative percentage of variance described as we move through the components. It shows how much of the total variance in the data is accounted for by each component and the collective outcome as we study multiple components. "Component 1" has the greatest eigenvalue (10.7312) in this analysis, accounting for 56.5 percent of the total variance. This shows that the first component essentially captures the variability of the data. Each component contributes less and less to the total variance explained as we go through the components.

Table 2.9 Total variance explained the water quality parameters in the Aik Stream.

Components	Initial Eigenvalue			Extraction sums of square loading		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
<b>COD</b>	5.702	30.012	30.012	5.702	30.012	30.012
<b>BOD</b>	3.339	17.575	47.588	3.339	17.575	47.588
<b>TDS</b>	1.663	8.751	56.338	1.663	8.751	56.338
<b>pH</b>	1.365	7.182	63.520	1.365	7.182	63.520
<b>T</b>	1.155	6.080	69.600	1.155	6.080	69.600
<b>NH<sub>3</sub>-N</b>	1.072	5.645	75.244	1.072	5.645	75.244
<b>TSS</b>	0.845	4.445	79.690			
<b>O &amp; G</b>	0.783	4.119	83.808			
<b>TOC</b>	0.522	2.746	86.555			
<b>NO<sub>3</sub>-N</b>	0.480	2.526	89.080			
<b>Cl-</b>	0.414	2.180	91.261			
<b>Cu</b>	0.363	1.913	93.174			
<b>Zn</b>	0.308	1.620	94.793			
<b>Cr</b>	0.255	1.343	96.136			
<b>Pb</b>	0.199	1.047	97.182			
<b>Cd</b>	0.174	.915	98.097			
<b>Ni</b>	0.144	0.760	98.858			
<b>As</b>	0.123	0.649	99.507			
<b>Hg</b>	0.094	0.493	100.000			



Table 2.10 The contribution of the six components of PCA in Aik Stream water quality.

<b>Component Matrix</b>							
	<b>Component</b>						
		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
HPZ	<b>COD</b>	0.720		0.286			
	<b>BOD</b>	0.532	0.525		0.239		
	<b>TDS</b>						
	<b>pH</b>						
	<b>T</b>	0.776					
	<b>NH<sub>3</sub>-N</b>	0.829					
	<b>TSS</b>	0.611					
	<b>O &amp; G</b>	0.801	0.309				
	<b>TOC</b>	0.774					
	<b>NO<sub>3</sub>-N</b>						0.354
	<b>Cl-</b>	0.573					
	<b>Cu</b>		0.831			0.385	
	<b>Zn</b>		0.870				
	<b>Cr</b>			0.667			
	<b>Pb</b>	0.703					
	<b>Cd</b>		0.460	0.502	0.559		
	<b>Ni</b>		0.414				
	<b>As</b>	0.845					
	<b>Hg</b>						0.832
MPZ		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
	<b>COD</b>						0.434
	<b>BOD</b>				0.439		
	<b>TDS</b>				0.203	0.777	
	<b>pH</b>		0.528				
	<b>T</b>	0.776					
	<b>NH<sub>3</sub>-N</b>	0.829					
	<b>TSS</b>	0.611					
	<b>O &amp; G</b>		0.309				
	<b>TOC</b>	0.774					
	<b>NO<sub>3</sub>-N</b>	0.242					0.354
	<b>Cl-</b>						
	<b>Cu</b>		0.831				
	<b>Zn</b>		0.870				
	<b>Cr</b>			0.567			
	<b>Pb</b>						
	<b>Cd</b>	-	0.460	0.502	0.559		
	<b>Ni</b>					0.487	
	<b>As</b>	0.845	0.392				
	<b>Hg</b>						
LPZ		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
	<b>COD</b>						
	<b>BOD</b>						
	<b>TDS</b>					0.577	
	<b>pH</b>	0.621	0.528				0.438



<b>T</b>						
<b>NH<sub>3</sub>-N</b>						
<b>TSS</b>		0.322				
<b>O &amp; G</b>						
<b>TOC</b>			0.444			
<b>NO<sub>3</sub>-N</b>	0.242			0.225		
<b>Cl-</b>						
<b>Cu</b>	0.499	0.631	0.432		0.385	
<b>Zn</b>						
<b>Cr</b>						
<b>Pb</b>						
<b>Cd</b>	0.668	0.560	0.502	0.559		
<b>Ni</b>					0.487	
<b>As</b>						
<b>Hg</b>						

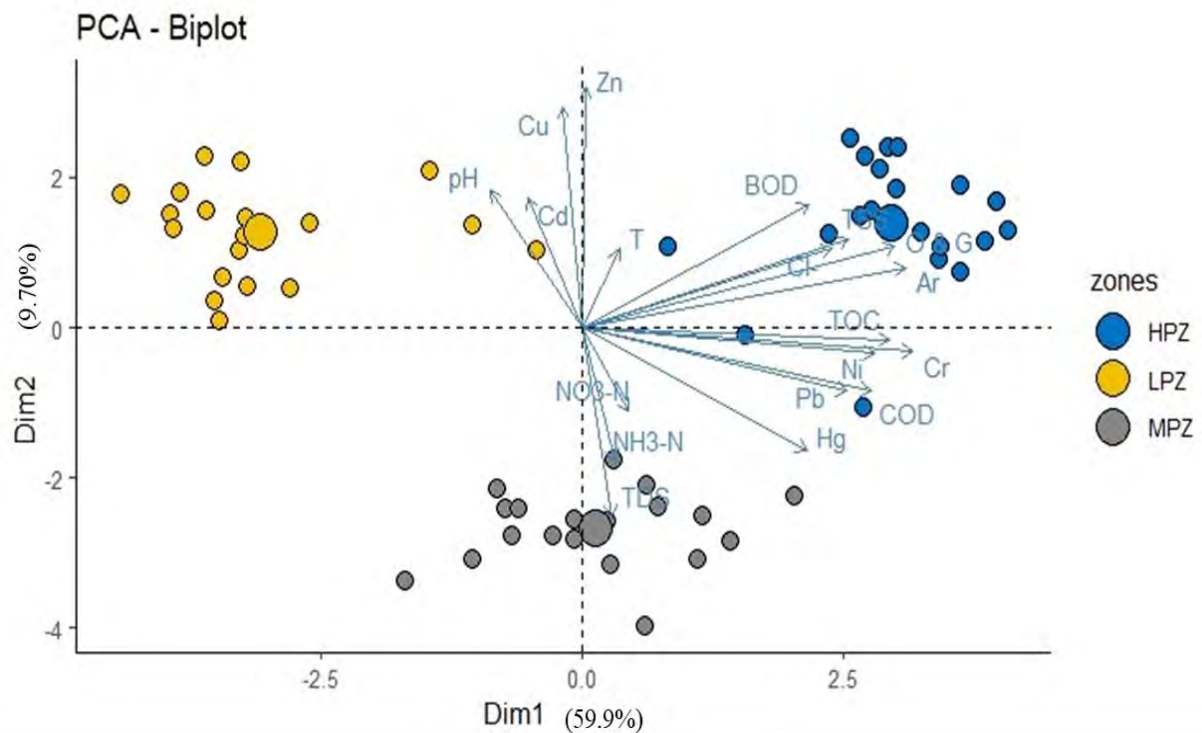


Figure 2. 12 PCA biplot depicting the three distant zones of the Aik stream

As shown in the biplot graphic derived from the PCA, the dataset is segregated into two dimensions, presented in Figure 2.11. PCA biplot has three different zones (HPZ, MPZ, and LPZ) that have been identified, and the biplot is used to show these zones in the context of the first two principal components, Dim1 and Dim2. These are two first and second principal components, respectively, obtained from PCA. This component indicates the linear and strong relationship among the original variables and PCA



components, capturing maximum data variation. Dim1 accounted for 59.9%, while Dim2 contributed 9.7% of the total variance. The zones (LPZ, MPZ, and HPZ) can identify the variable grouping in the dataset with similar characteristics. The zone is based on the score of variables in the dataset analog with PCA. The HPZ represents the group of variables with high values on Dim1, the LPZ indicates the low values, and MPZ may be falling in between. Fig 2.12. The Scree plot of PCA shows a sharp decline in eigenvalues after the first six components. This suggests that retaining these initial six components would be optimal for capturing the majority of variance in water quality parameters while also preventing overfitting.

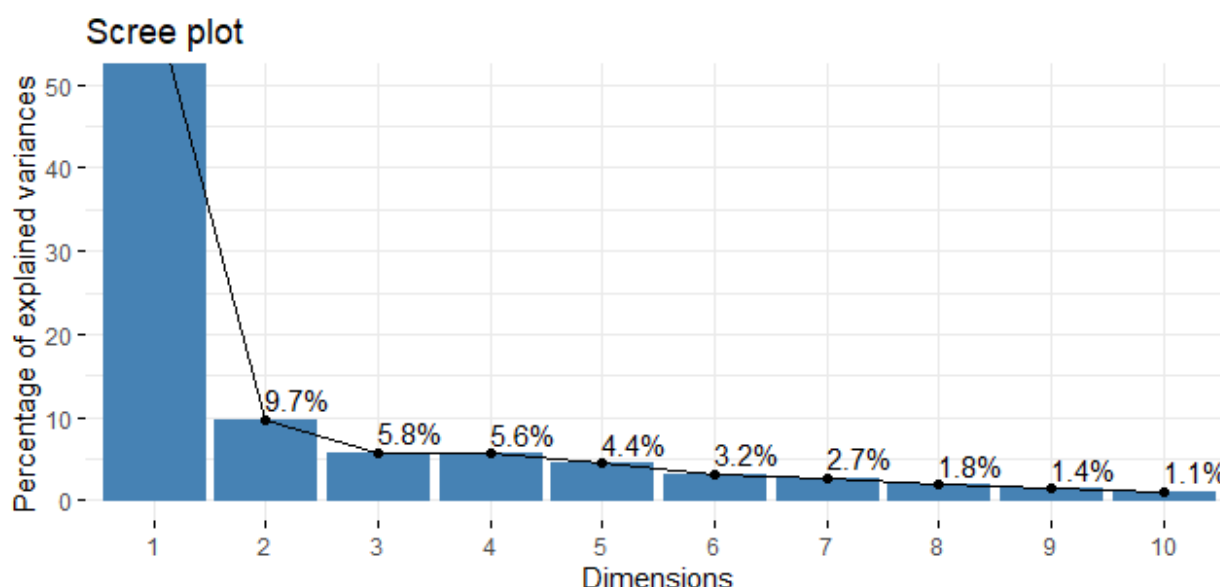


Figure 2. 13 Scree plot for PCA component selection

### 2.3.4 Innovative Monitoring with Machine Learning Models

#### 2.3.4.1 Calculation of WQI

The WQI was calculated using the CCME Calculator Software Version 1.0 given by the Canadian Council of Ministers of the Environment (Table 2.11 and Appendix 2). The observed WQI demonstrated that the upstream portion with 51 monitoring sites was of good and fair quality ( $64 < I \leq 94$ ), while the midstream with 66 monitoring sites was mainly classified as poor water quality ( $0 < I \leq 44$ ). The downstream portion with 33 monitoring sites had marginal values ( $45 < I \leq 64$ ). The water quality in the upstream region is comparatively better and exhibits minimal contamination, primarily due to the absence of



industrialization (NN & NN, pers. obs.). Nevertheless, industrial and sewage waste pollutants gradually accumulate once they traverse the city, resulting in a substantial build-up of various contaminants, including heavy metals. Thus, the water at the mid-stream undergoes excessive pollution, threatening its overall quality. Moving downstream, the water becomes moderately polluted due to a decline in industrial activity in that area. The deterioration of water quality from upstream to downstream of Aik-stream is mainly linked to urban and industrial wastewater discharges in that area.

### **2.3.5 Evaluating ML models**

ML models vis., Gradient Boosting, and Random Forests were employed on a dataset comprising 150 observations for 19 water quality parameters to predict the Water Quality Index (WQI) of the Aik-Stream. To mitigate overfitting, the data was divided into two distinct subsets: the first, encompassing 80% of the data, was designated for the training phase, allowing us to grasp the system's dynamics thoroughly. The second subset, comprising 20% of the data, enables us to evaluate and verify the performance of the predictive models. The development of these predictive models involved using the feature elimination-linear algorithm RFE-L, which helped identify a pertinent subset of features from a pool of nineteen potential input variable combinations (1–19). To evaluate the predictive performance of these models for the Water Quality Index (WQI), we employed a set of statistical metrics, including  $R^2$ , RMSE, MAE, RAE, and RRSE, during both the training and testing phases (Table 2.12 and 2.13)

Table 2.11 compares the GB model's predictive performance during the training and testing phases. The GB-4th input combination (with training  $R^2 = 0.94$ , training RMSE = 6.33, testing  $R^2 = 0.90$ , testing RMSE = 6.54) demonstrated the most optimal performance among the various GB models examined in this study. It is important to note that predictive performance typically leans more favorably towards the training phase as compared to the testing phase, which is an expected outcome during the model training process aimed at minimizing predictive errors. The GB model is initially constructed during the training stage and subsequently assessed during the testing phase. Furthermore, the GB–13th input combination also exhibited better predictive performance in the training ( $R^2=0.88$ , RMSE= 7.24) and testing ( $R^2=0.85$ , RMSE = 8.67) stages than other input combination models. Although the GB–4<sup>th</sup> input combination model, consisting of sixteen input parameters (predictors), offers the best predictive performance, the GB–13th input combination model, which incorporates only seven input parameters, namely COD, TOC, OG,  $\text{NH}_3\text{N}$ , As, Ni, and



Zn, also provides satisfactory performance ( $R^2$  train=0.88,  $R^2$  test=5.55). Therefore, the GB–13th input combination model is identified as the superior predictive model, requiring a more limited number of input physicochemical variables. Overall, the results demonstrate that the GB model delivers high predictive accuracy for water quality in both the training and testing stages.

Moving on to (Table 2.12), it highlights the predictive performance of the RF model. The RF-2nd input combination has remarkable predictive capabilities in the training phase ( $R^2 = 0.93$ , RMSE = 6.33) and the testing phase ( $R^2 = 0.85$ , RMSE = 8.02). However, the 10th input combination, despite utilizing a smaller number of variables, also delivers commendable results in both training ( $R^2=0.85$ , RMSE=8.67) and testing ( $R^2=0.84$ , RMSE=9.34).

Table 2.11 Evaluation Measures for the GB Algorithm in WQI Prediction for the Aik-Stream (Training and Testing Datasets).

MODEL	$R^2$	RMSE	MAE	RAE (%)	RRSE (%)
TRAINING					
1	0.93	6.45	9.35	82.23	21.57
2	0.79	7.32	5.37	99.45	53.22
3	0.75	4.38	10.81	109.11	62.82
4	<b>0.94</b>	<b>6.33</b>	<b>4.62</b>	<b>38.21</b>	<b>39.71</b>
5	0.82	6.47	11.12	70.55	58.67
6	0.75	7.37	6.25	91.32	50.09
7	0.75	4.46	7.35	53.25	87.91
8	0.73	7.56	11.45	44.45	51.81
9	0.83	9.56	13.21	35.65	89.22
10	0.81	8.26	12.23	62.12	91.45
11	0.80	8.67	12.56	110.23	83.21
12	0.77	6.61	7.41	37.23	72.36
13	<b>0.88</b>	<b>7.24</b>	<b>10.15</b>	<b>16.34</b>	<b>73.56</b>
14	0.72	5.23	5.85	99.23	83.56
15	0.83	4.45	7.57	98.45	64.34
16	0.87	4.2	12.32	71.67	104.67
17	0.71	7.34	4.23	98.89	96.67
18	0.78	8.35	7.23	76.34	67.34
19	0.85	7.36	4.23	101.23	101.56
TESTING					
1	0.91	7.05	9.15	87.21	27.01
2	0.84	11.34	13.63	83.23	118.56
3	0.77	10.76	8.53	48.54	85.32
4	<b>0.90</b>	<b>6.54</b>	<b>8.65</b>	<b>59.32</b>	<b>55.32</b>
5	0.79	11.43	7.54	50.23	77.32
6	0.66	10.32	13.76	37.31	86.32
7	0.74	8.31	12.98	38.41	65.32
8	0.70	8.54	11.76	90.12	60.34
9	0.81	7.43	12.32	48.23	86.12
10	0.77	10.25	13.32	57.67	53.23
11	0.79	10.21	7.31	59.34	52.45
12	0.72	11.24	12.41	38.65	72.67
13	<b>0.85</b>	<b>8.67</b>	<b>11.31</b>	<b>19.31</b>	<b>8.32</b>



14	0.84	7.54	8.31	42.45	110.56
15	0.85	10.25	8.21	118.12	113.78
16	0.64	12.43	7.41	119.23	112.54
17	0.72	8.67	9.32	81.21	85.23
18	0.75	6.78	9.31	58.12	57.3
19	0.73	8.67	11.31	41.12	109.23

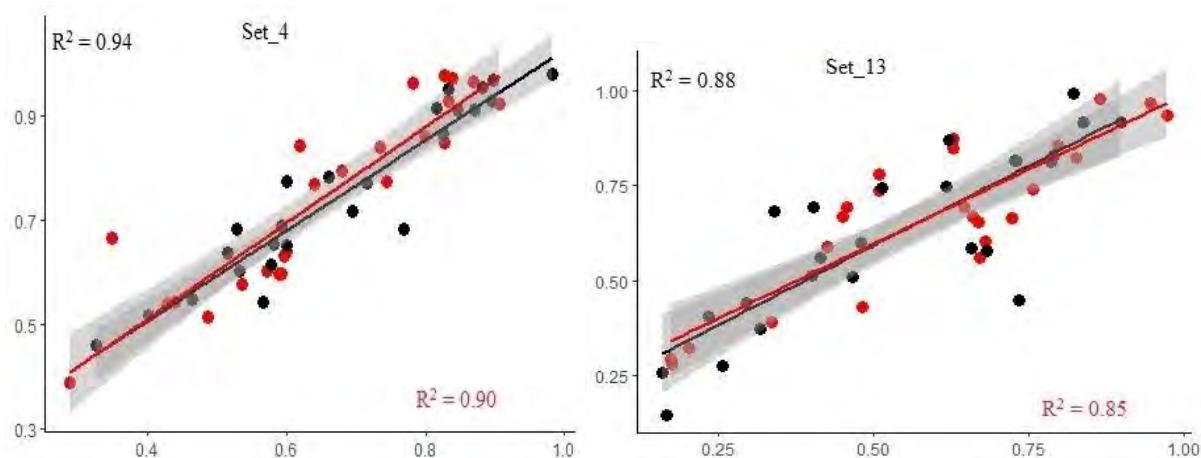


Figure 2. 14 The top-performing input variables (4th and 13th) for WQI Prediction, with training data shown in black and testing data in red.

Table 2.12 Evaluation Measures for the RF Algorithm in WQI Prediction for the Aik-Stream (Training and Testing Datasets).

MODEL	R <sup>2</sup>	RMSE	MAE	RAE (%)	RRSE (%)
TRAINING					
1	0.89	7.45	10.34	89.45	99.76
2	<b>0.93</b>	<b>5.23</b>	<b>6.45</b>	<b>32.84</b>	<b>55.33</b>
3	0.78	8.45	11.56	75.54	64.32
4	0.63	8.45	14.67	25.23	31.76
5	0.71	8.45	14.56	52.43	72.23
6	0.73	13.78	10.81	96.45	94.24
7	0.87	6.9	10.89	70.56	60.62
8	0.61	12.27	13.32	92.23	85.23
9	0.78	8.67	13.76	108.34	103.98
10	<b>0.85</b>	<b>8.67</b>	<b>9.23</b>	<b>35.21</b>	<b>72.34</b>
11	0.75	12.67	6.57	56.56	75.12
12	0.79	8.78	12.87	18.56	83.23
13	0.88	9.34	14.52	86.56	80.34
14	0.70	5.67	11.51	113.45	65.45
15	0.77	6.98	12.65	86.78	116.23
16	0.79	8.45	12.54	66.56	65.56
17	0.67	13.67	8.67	74.32	79.12
18	0.89	7.67	14.23	79.45	74.23
19	0.87	7.98	12.26	88.12	60.09
TESTING					
1	0.75	10.92	7.12	96.41	81.21
2	<b>0.91</b>	<b>8.02</b>	<b>9.24</b>	<b>51.34</b>	<b>75.34</b>
3	0.74	7.23	10.91	97.21	51.45



4	0.82	11.34	10.98	48.19	86.45
5	0.77	12.61	8.56	80.81	78.87
6	0.83	11.34	6.54	64.72	67.34
7	0.84	9.45	12.45	98.63	64.32
8	0.71	10.45	11.63	62.53	67.35
9	0.75	10.45	12.45	93.49	74.45
10	<b>0.84</b>	<b>9.34</b>	<b>10.76</b>	<b>58.61</b>	<b>64.23</b>
11	0.73	9.45	8.47	85.54	85.45
12	0.74	9.45	13.87	71.51	70.45
13	0.76	8.87	12.65	78.34	81.56
14	0.65	9.65	11.73	121.61	58.31
15	0.75	10.65	6.63	81.61	83.75
16	0.75	10.43	11.52	92.71	60.56
17	0.64	5.89	12.61	84.87	88.23
18	0.84	7.62	7.52	64.23	76.67
19	0.67	6.71	10.76	48	85.23

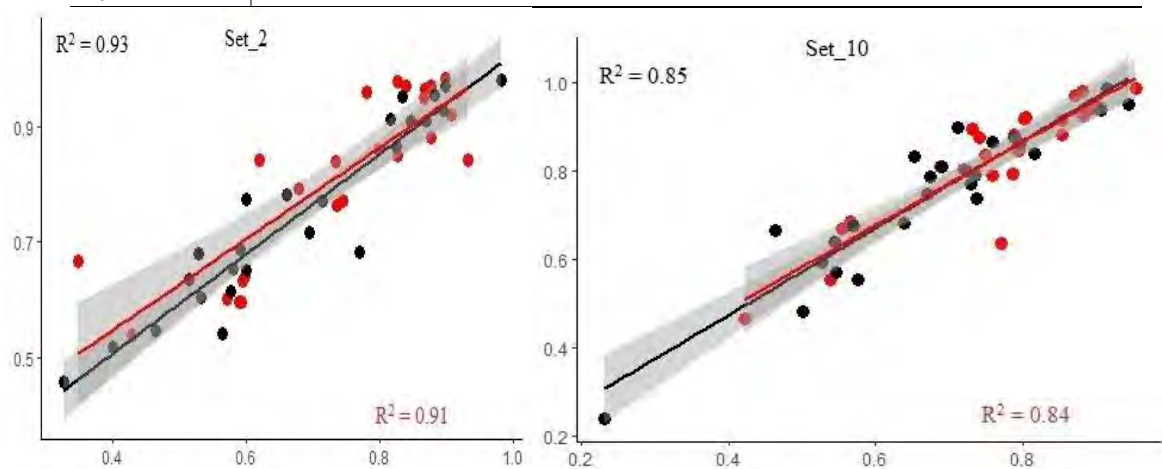


Figure 2. 15 RF Algorithm's Best Inputs (2nd and 10th) for WQI Prediction (Training: Black, Testing: Red).

The overall comparison analysis of the ML models showed that both models demonstrated robust performance. In contrast, the Gradient Boost model distinguished itself from RF model with superior predictive abilities, and its (GB–13th) input combination model is depicted to be the better predictive model, which needs a limited number of input physicochemical variables.

### 2.3.6 Identifying Optimal Input Combinations

Best subset regression analysis was done to ascertain our WQ model's most effective input combinations. To achieve this, we computed six statistical criteria, including Mean Squared Error (MSE), determination coefficients ( $R^2$ ), adjusted  $R^2$ , Mallows' Cp (Gilmour 1996), Akaike's AIC, and BIC. The outcomes of these calculations are displayed in (Table 2.3.12), where it becomes evident that subset 15 stands out as the preferred option among the various models. It exhibited the lowest MSE (4.034), the lowest AIC (351.54), the lowest BIC (579.42), the lowest Mallows' Cp (10.54), and the highest  $R^2$  (0.93) and adjusted  $R^2$  (0.92).



As a result, subset 15 was identified as the most suitable input combination for predicting the WQI model table 2.14).

Moreover, to assess the potential multicollinearity among the water quality index parameters, we utilized the Variance Inflation Factor (VIF) and its reciprocal,  $1/VIF$ , as presented in (Table 2.14). A widely accepted criterion is that  $1/VIF$  should be less than 0.1, and VIF should be less than 10 to indicate the absence of multicollinearity about the target variable, WQI. The findings in (Table 2.14) reveal that the majority of variables met the criteria with  $1/VIF$  values below 0.1. However, during the VIF assessment,  $NH_3N$  exhibited a value of 10.32, and As had a value of 11.43, slightly surpassing the VIF threshold 10. Nevertheless, these results do not suggest the presence of multicollinearity.

Table 2.13 Displays the results of a regression analysis aimed at identifying optimal input combinations for modelling WQI.

No of parameters	Name of parameters	MSE	R <sup>2</sup>	Adj R <sup>2</sup>	AIC	BIC	Malkow PC
1	TDS	15.34	0.83	0.73	331.64	551.02	54.32
2	TDS, TSS	13.67	0.88	0.77	334.45	555.11	24.98
3	TDS, TSS, Pb	8.870	0.82	0.80	337.01	557.09	21.94
4	TDS, TSS, Pb, pH	11.56	0.82	0.81	337.67	559.34	12.09
5	TDS, TSS, Pb, pH, T	12.43	0.85	0.83	339.45	560.04	9.099
6	TDS, TSS, Pb, pH, T, Hg	14.83	0.85	0.83	340.69	562.87	11.32
7	TDS, TSS, Pb, pH, T, Hg, NO <sub>3</sub> N	7.450	0.82	0.80	342.81	563.82	15.98
8	TDS, TSS, Pb, pH, T, Hg, NO <sub>3</sub> N, BOD	7.321	0.87	0.88	344.91	564.82	23.12
9	TDS, TSS, Pb, pH, T, Hg, NO <sub>3</sub> N, BOD, Cl	7.221	0.84	0.83	345.21	567.45	22.87
10	TDS, TSS, Pb, pH, T, Hg, NO <sub>3</sub> N, BOD, Cl, Cr	6.898	0.84	0.83	346.41	568.32	21.54
11	TDS, TSS, Pb, pH, T, Hg, NO <sub>3</sub> N, BOD, Cl, Cr, Cd	6.854	0.88	0.87	347.23	569.22	32.54
12	TDS, TSS, Pb, pH, T, Hg, NO <sub>3</sub> N, BOD, Cl, Cr, Cd, COD	7.342	0.88	0.86	347.88	572.67	11.54
13	TDS, TSS, Pb, pH, T, Hg, NO <sub>3</sub> N, BOD, Cl, Cr, Cd, COD, TOC	6.342	0.89	0.88	348.11	574.44	12.45
14	TDS, TSS, Pb, pH, T, Hg, NO <sub>3</sub> N, BOD, Cl, Cr, Cd, COD, TOC, NH <sub>3</sub> N	6.811	0.89	0.87	349.03	574.81	11.32
15	<b>TDS, TSS, Pb, pH, T, Hg, NO<sub>3</sub>N, BOD, Cl, Cr, Cd, COD, TOC, NH<sub>3</sub>N, Zn</b>	<b>4.034</b>	<b>0.93</b>	<b>0.92</b>	<b>351.54</b>	<b>579.42</b>	<b>10.54</b>
16	TDS, TSS, Pb, pH, T, Hg, NO <sub>3</sub> N, BOD, Cl, Cr, Cd, COD, TOC, NH <sub>3</sub> N, Zn, As	6.231	0.90	0.87	352.22	576.43	18.87
17	TDS, TSS, Pb, pH, T, Hg, NO <sub>3</sub> N, BOD, Cl, Cr, Cd, COD, TOC, NH <sub>3</sub> N, Zn, As, Ni	7.432	0.89	0.87	353.87	580.42	17.97



<b>18</b>	TDS, TSS, Pb, pH, T, Hg, NO <sub>3</sub> N, BOD, Cl, Cr, Cd, COD, TOC, NH <sub>3</sub> N, Zn, As, Ni, Cu	5.213	0.88	0.84	353.81	581.33	15.74
<b>19</b>	TDS, TSS, Pb, pH, T, Hg, NO <sub>3</sub> N, BOD, Cl, Cr, Cd, COD, TOC, NH <sub>3</sub> N, Zn, Ni, As, Cu, OG	5.321	0.90	0.83	354.31	583.99	11.53

Table 2.14 Multicollinearity statistics analysis for WQI parameters

Variable	VIF	1/VIF
COD	7.34	0.01239
TOC	9.47	0.042611
OG	9.15	0.06599
NH <sub>3</sub> N	10.32	0.069834
As	11.43	0.087453
Ni	8.41	0.011895
Zn	7.98	0.0125269
Cd	6.47	0.0154488
Cr	5.95	0.0168013
Cl	4.97	0.01308
BOD	4.36	0.029166
TDS	4.32	0.031383
TSS	4.31	0.031921
Pb	4.2	0.037852
pH	3.22	0.010983
T	3.2	0.012555
Hg	2.26	0.043095
NO <sub>3</sub> N	1.93	0.018168
Cu	1.79	0.059737

## 2.4 Discussion

Aik Stream is a carrier of wastewater that receives municipal sewage and industrial effluents before flowing into the Chenab River. Industries along the Aik stream in the Sialkot metropolitan area do not follow the national criteria for sanitation and discharge their wastewater without first treating it (A. Qadir et al., 2008a). The high levels of hazardous materials in the mid-stream area also suggest that the relevant authorities are not complying with regulations for properly disposing of domestic and commercial wastewater, which contaminates the water. According to studies (Iqbal, Shoaib, Agwanda, & Lee, 2018; Ranjith, 2019; Revankar & Kadadevaru, 2021) has been highlighted the potential connection between surface water pollution in streams and rivers in India and Pakistan and the untreated release of industrial and municipal waste. Industrial and urban wastewater contains organic and inorganic components that utilize dissolved oxygen, reducing oxygen levels (Mokarram,



Saber, & Sheykhi, 2020). Over the past thirty years, exports of tanned leather from Pakistan have increased while the demand for the product in the developed world has declined due to stricter environmental laws. Since the Pakistan Environmental Protection Agency has not strengthened its regulation, the unchecked growth of this business has caused significant damage to the ecosystem. Untreated wastewater from industry and municipalities is carelessly discharged into surrounding ditches, ponds, streams, rivers, and open spaces.

Tanning processes generate large amounts of wastewater containing Cr, Pb,  $\text{Cl}^-$ , salts, and organic nutrients (China et al., 2020). The high concentration of these hazardous compounds in the study area was associated with the widespread usage of  $\text{Na}_2\text{S}$  and Cr salts in chrome tanning. Assessment of water quality at mid-stream sites for heavy metals such as chromium (Cr), lead (Pb), copper (Cu), arsenic (As), nickel (Ni), mercury (Hg), zinc (Zn), and cadmium (Cd) was observed that their concentrations exceeded the guidelines set by the World Health Organization (Organization, 2021). It is well known that releasing these metals into surface water could have toxicological consequences (Khadija et al., 2021).

The high Cr concentration found in this study is mainly due to wastewater discharged into streams by tanneries and electroplating companies, which may be deposited in higher concentrations in the sediment (Nur-E-Alam, Mia, Ahmad, & Rahman, 2020). The maximum Cr concentration was reported at the sites located near Sialkot city. The result showed that high amounts of chromium (Cr) were recorded in HPZ (1.981 mg/L) and MPZ (1.172 mg/L), which is due to the significant pollution from tanneries. The chromium concentration (Cr) in HPZ and MPZ exceeded the permissible 0.05 mg/L set by the World Health Organization (Organization 2020). The tanning process produces a significant amount of wastewater that contains high concentrations of chromium and soluble salts such as sodium chloride ( $\text{NaCl}$ ) (Moten & Sami, 2000). The indiscriminate use of chromium salts in tanneries is one of the main reasons for increased chromium content in the wastewater and streams of Sialkot. Chromium (Cr) is predominantly contained in the industrial waste resulting from chrome tanning, with about 70% being absorbed by the hides, while the remaining 30% is not absorbed and instead ends up in wastewater and sludge. It has been found that sustained release of Cr, even at low concentrations, can harm aquatic life and disrupt the food chain of aquatic ecosystems (Vaipoulou & Gikas, 2020). A high pH allows Cr to transform into complex substances and become part of suspended particulate matter that settles as the wastewater travels the distance from the source. Similar to previous reports (R. Nazir et al., 2015; Sankhla, Kumar, & Prasad, 2019), the concentration of the current study was higher.



Industrial and urban wastewater are the primary sources responsible for increased nickel (Ni) content in stream water (V. Kumar, Parihar, et al., 2019). The highest nickel (Ni) concentration in suspended particulate matter was observed, possibly due to the discharge of municipal sewage and industrial wastewater containing waste from electroplating plants used to manufacture surgical instruments. In an alkaline pH environment, dissolved nickel (Ni) metal reacts with suspended particles and organic substances, resulting in their deposition in the stream bed. In the current study, the concentration of Ni was found to be higher than the concentrations reported in studies by (Delina et al., 2020; M. S. Islam, Ahmed, Raknuzzaman, Habibullah-Al-Mamun, & Islam, 2015).

Pb in dissolved form was recorded at its maximum in HPZ (2.015) and lowest in LPZ (0.86 mg/L). Pb concentration increases when moved from the upstream site to midstream sites. Other sources of lead (Pb) include manufacturing processes such as battery factories, as well as activities related to paints and pigments, as well as the incineration of municipal and hazardous waste (Bouzekri, El Fadili, El Hachimi, El Mahi, & Lotfi, 2020). Lead (Pb) comes from vehicle exhaust and is deposited from the atmosphere onto the ground, especially near major highways (Frank, Poulakos, Tornero-Velez, & Xue, 2019). The concentration of lead (Pb) in the water of Aik-Stream exceeded the acceptable threshold of 0.01 mg/L for freshwater and was higher compared to the results in the studies of (Kennedy, Koth, & Carruth, 2015; V. Kumar, Parihar, et al., 2019). (Mokarram et al., 2020) they claimed that various factories were responsible for the elevated concentrations of zinc (Zn), copper (Cu), nickel (Ni), iron (Fe), aluminum (Al), lead (Pb), manganese (Mn), chromium (Cr), and tin (Sn), which leads to damage to the water quality of the river. The National Water Services Commission (NWSC) has revealed that the release of industrial wastewater is a factor contributing to increasing water pollution in rivers (Chen, Wang, & Yao, 2023). Furthermore, according to the Selangor Water Management Authority (SWMA), industrial activity was the reason for an increase in ammonia concentration ( $> 0.7$  mg/L) in water between 2012 and 2015 (Farid, Lubna, Choo, Rahim, & Mazlin, 2016; Mokarram et al., 2020). In recent years, multivariate statistical analysis has become one of the most used methods for assessing water quality and conducting analyses related to environmental pollution. Principal Component Analysis (PCA) was applied on the Songkhram River, Thailand (Muangthong & Shrestha, 2015) and on the Jia Bharali River basin, India, to estimate the number of factors to reduce assess water quality (Khound & Bhattacharyya, 2017). (K. P. Singh, Malik, Sinha, Singh, & Murthy, 2005) used 24 different parameters to classify the pollution levels of Gomti River



water into three categories: low, moderate, and high. Their results showed that Principal Component Analysis (PCA) can enhance the reliability of statistical analyses by reducing the number of parameters used in water quality assessment. (S. Shrestha et al., 2007) used multivariate statistical analysis to assess the water quality of Japan's Fuji River. They also classified the level of river pollution into three categories: low, moderate, and high pollution. Their report highlighted that parameters such as Biological Oxygen Demand (BOD), pH, Electrical Conductivity (EC), nitrate ( $\text{NO}_3^-$ ), and ammonium ( $\text{NH}_4^+$ ), have the greatest impact on water quality. (Mustapha et al., 2014) used Analysis of Variance (ANOVA) approach to assess the spatial variations in surface water quality at the upper reaches of the Kano River in Nigeria. An analysis of surface water quality fluctuation in the Xiangjiang River, China, was conducted using ANOVA in a companion study (Z. Zhang et al., 2010).

The water quality index (WQI) is a lumped index based on concentrations of different water quality constituents (Poonam, Tanushree, & Sukalyan, 2013), which managers could use to assess the overall quality of various water resources (Lumb, Sharma, & Bibeault, 2011). (Ramakrishnaiah, Sadashivaiah, & Ranganna, 2009) utilized the WQI index to assess water quality in India. They incorporated 12 parameters of pH, total hardness (TH), calcium ( $\text{Ca}^{2+}$ ), magnesium ( $\text{Mg}^{2+}$ ), bicarbonate, chloride ( $\text{Cl}^-$ ), nitrate ( $\text{NO}_3^-$ ), sulfate, total dissolved solids (TDS), iron ( $\text{Fe}^{2+}$ ), manganese ( $\text{Mn}^{2+}$ ), and fluoride ( $\text{F}^-$ ).

Various methods exist for assessing surface water quality that are time-efficient and cost-effective. Machine learning approaches have been shown to enhance efficiency and reduce the costs of assessing surface water quality. Research findings suggest that machine learning techniques effectively minimize the necessary parameters and monitor stations for water quality evaluation (Alzubi, Nayyar, & Kumar, 2018). Recently, a growing interest has been in using machine-learning models to assess water quality. These models have the potential to quickly and effectively detect changes in water quality conditions. Our study aimed to contribute to this field by evaluating different machine learning models for predicting water quality. Many researchers have used the CCME-WQI as a crucial variable in their machine-learning models to predict water quality trends. This index has also been used in other studies to forecast water quality (Yilma, Kiflie, Windsperger, & Gessese, 2018). Our findings align with a study by (Khoi, Quan, Linh, Nhi, & Thuy, 2022), which found that boosting-based algorithms, especially Extreme Gradient Boosting (XGBoost), were highly accurate in predicting water quality for the La Buong River in Vietnam. With an R-squared value of 0.989 and a low RMSE (Root Mean Square Error) of 0.107, XGBoost proved its



effectiveness in this specific context. Other notable research, such as the work of (Asadollah, Sharafati, Motta, & Yaseen, 2021), highlighted the superiority of the Extra Tree Regression (ETR) model. In their study, ETR outperformed other models with a high R-squared value of 0.97, demonstrating its effectiveness in predicting water quality when considering ten variables.

Researchers in various countries have employed different models with R-squared values exceeding 0.90, including (J. Li et al., 2019) in Iraq (Kamyab-Talesh, Mousavi, Khaledian, Yousefi-Falakdehi, & Norouzi-Masir, 2019) in Iran, (Nathan, Saravanane, & Sundararajan, 2017) in India, and (Gazzaz, Yusoff, Aris, Juahir, & Ramli, 2012) in Malaysia. A comparative analysis by (Singha, Pasupuleti, Singha, Singh, & Kumar, 2021) ranked XGBoost as the top-performing model, followed by Artificial Neural Networks (ANN) and Random Forest (RF). Other studies, such as the research conducted by (Sakaa et al., 2022), consistently favored Random Forest over Support Vector Regression (SVR) in predicting Water Quality Index (WQI). However, Support Vector Regression (SVR) performed better than Artificial Neural Networks (ANN) in the study conducted by (Hazarika, Gupta, Ashu, & Berlin, 2020).

Similarly, (Wong et al., 2022) conducted an extensive comparison of five regression models, including Multilayer Perception (MLP), Random Forest (RF), Decision Tree Regression (DTR), AdaBoost, and Support Vector Regression (SVR). Their results showed that the Random Forest algorithm outperformed the others with a high R-squared value of 0.974. Besides ensemble methods like Random Forest and XGBoost, alternative machine learning algorithms, such as Bayesian Regularization (Sakizadeh, 2015) and Artificial Neural Networks (ANN) by (Gazzaz et al., 2012), have shown strong correlations between predicted and observed water quality values in previous research. Comparative analyses among other ML models, different from the ones we mentioned, generally favored the Decision Tree model over algorithms like Naive Bayes (NB), k-nearest Neighbor (KNN), Multilayer Perceptron (MLP), and Logistic Regression (LogR) (Ahmed, Mumtaz, & Hassan Zaidi, 2021). Feature importance analysis by (Wong et al., 2022) showed that the proposed modified RF model, which included relatively important novel variables, is more proficient in water quality modeling. The predictive performance of the SVM model was observed to be reduced when irrelevant parameters were included in the input dataset (Leong, Bahadori, Zhang, & Ahmad, 2021). Similar findings were obtained by (Gazzaz et al., 2012; Sakaa et al., 2022). Therefore, it is noteworthy to select carefully and include only the relevant parameters



to maximize the benefits of ML models in prediction tasks. Our research highlights that Gradient Boost and Random Forest models are the preferred machine learning models for achieving accurate predictive results, surpassing other prediction models found in existing literature. These models have proven their usefulness in various geographic contexts, providing valuable insights for water quality management and monitoring. Researchers are continuously investigating novel algorithms and techniques to improve the accuracy and resilience of machine learning models.

The results of this study show that industrial processes are influencing the deterioration of stream water quality, either directly or indirectly. Evidence of diffuse source water contamination was found in areas designated as Low Pollution Zones (LPZ). High Pollution Zones (HPZs) were regions with markedly declining water quality due to heavy pollution from specific sources, such as tanneries, industrial waste, and municipal sewage from Sialkot City. Sites further downstream that were impacted by non-point source pollution and household waste from nearby smaller towns were included in the Medium Pollution Zones (MPZ). Rainfall during the monsoon season causes floods in the Aik-Stream, which causes harmful contaminants, especially heavy metals, to accumulate in the nearby areas. The hundreds of water pumps along the Aik-Stream are also used for irrigation. (K. P. Singh et al., 2005). Moreover, (A. Qadir et al., 2008a) noted that directly taking water for agriculture from rivers and tiny streams is common in Pakistan and India. If this keeps up, there is a real chance that agricultural soils may get seriously contaminated with metals, which could damage crops even more. Municipal and industrial garbage is increasingly finding their way into the Aik-stream. The loss of plant species and different biological communities from the aquatic environment could result from such deterioration, which could make the ecosystem of the Aik-stream even more fragile.

This study sheds light on the adverse consequences of contamination in the Aik-stream on neighboring soil ecosystems. As the stream increasingly becomes a channel for municipal and industrial waste, it facilitates the migration of numerous pollutants to adjacent land areas. This migration primarily occurs via leaching and precipitation, a phenomenon more pronounced during monsoon spells. In these times, heightened flow in the Aik-stream causes contaminated water to spread across the land, seeping pollutants like heavy metals and other industrial residues into the nearby soils. This infiltration changes the soil composition and might disturb its natural biogeochemical balance. The ramifications of such seepage are manifold. The altered soil properties can jeopardize its fertility and overall health. The



building up of these pollutants also introduces a notable threat to plants, potentially hampering their growth and physiological functions. The uptake of these pollutants by plant roots brings forth issues related to bioaccumulation and the potential for these toxins to enter terrestrial and aquatic food chains.

Moreover, the study emphasizes the potential risks to human health from consuming crops watered with this contaminated runoff, highlighting broader concerns about environmental well-being and food security. The insights from this investigation make a strong case for the initiation of robust waste management and pollution control strategies along the Aik stream to curb its adverse ecological and health repercussions. Such efforts require a holistic approach, merging environmental stewardship, policy development, and community involvement to preserve the vitality of these ecosystems. It is imperative to reconsider the prolonged use of such water for irrigation. Persistently irrigating water with elevated levels of heavy metals can lead to an increased presence in the soil, potentially affecting plant functions and posing significant risks to human health. (M. A. Khan & Ghouri, 2011).



## 2.5 Conclusion

This chapter has unveiled significant degradation in Aik-Stream's water quality, notably in its mid to downstream parts, primarily stemming from extensive polluting activities in the catchment area. Industrial operations have profoundly altered the stream into an environmentally precarious and unacceptable state. Spatial analysis categorized the stream into three distinct zones. The first is the Less Polluted Zone (LPZ), which has relatively better water quality. The High Polluted Zone (HPZ) displays the most severe degradation, and the Moderately Polluted Zone (MPZ) has less impact on water quality. Multivariate techniques effectively identified vital factors contributing to this degradation, pinpointing their likely sources. The findings underscored that critical water quality parameters, including COD, BOD, TDS,  $\text{NO}_3\text{-N}$ ,  $\text{NH}_3\text{-N}$ ,  $\text{Cl}^-$ , and various heavy metals (Cu, Zn, Cd, As, Cr, Hg, Ni, and Pb), consistently exceeded permissible levels, particularly in the midstream section of Aik-Stream. In addition to traditional techniques, this study explored a novel approach by incorporating machine learning models for efficient water quality index (WQI) prediction with fewer variables and reduced time. Two machine learning models, GB (Gradient Boosting) and RF (Random Forests), were used to forecast the water quality index of Aik-Stream. Our findings indicate that both models excel in predicting the water quality index, with GB slightly trailing the RF model in predictive accuracy.

Nevertheless, they both offer reasonably accurate predictions. The GB model achieves its highest predictive accuracy using only seven input variables: COD, TOC, OG,  $\text{NH}_3\text{N}$ , Ar, Ni, and Zn. This discovery suggests that the GB model, particularly with the 13th input combination, has the potential to empower water managers and policymakers to calculate the water quality index for rivers and streams efficiently. This approach can be achieved with reduced computational time, lower costs, and less real-time monitoring at multiple polluted sites, enhancing water resource management strategies and leading to more effective water quality assessments for sustainable river water management. Additionally, this study highlights the adverse effects of Aik-stream contamination on soil ecosystems, emphasizing pollutant migration and its risks to plant growth and human health, underscoring the urgent need for robust waste management and pollution control strategies.





Contents lists available at ScienceDirect

## Journal of Cleaner Production

journal homepage: [www.elsevier.com/locate/jclepro](http://www.elsevier.com/locate/jclepro)



# Monitoring the Industrial waste polluted stream - Integrated analytics and machine learning for water quality index assessment

Ujala Ejaz<sup>a, b</sup>, Shujaul Mulk Khan<sup>a, c</sup>, Sadia Jehangir<sup>a, b</sup>, Zeeshan Ahmad<sup>c, \*\*\*</sup>,  
Abdullah Abdullah<sup>a</sup>, Majid Iqbal<sup>a, d</sup>, Noreen Khalid<sup>e</sup>, Aisha Nazir<sup>f</sup>, Jens-Christian Svenning<sup>b, \*\*</sup>

<sup>a</sup> Department of Plant Sciences, Quaid-i-Azam University, Islamabad, 45320, Pakistan

<sup>b</sup> Center for Ecological Dynamics in a Novel Biosphere (ECONOVO), Department of Biology, Aarhus University, Ny Munkegade 114, DK-8000, Aarhus C, Denmark

<sup>c</sup> CAS Key Laboratory of Tropical Forest Ecology, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengli, 665303, China

<sup>d</sup> Institute of Geographic Sciences and Natural Resources Research, University of Chinese Academy of Sciences, China

<sup>e</sup> Department of Botany, Government College Women University, Sialkot, Pakistan

<sup>f</sup> Environmental Biotechnology Research Laboratory, Institute of Botany, University of the Punjab, Lahore, Pakistan

## ARTICLE INFO

Handling editor: Jin-Kuk Kim

### Keywords:

Industrial wastewater

Machine learning

Assessment and monitoring

Water Quality Index

Artificial Intelligence

## ABSTRACT

The Water Quality Index (WQI) is a primary metric used to evaluate and categorize surface water quality which plays a crucial role in the management of fresh water resources. Machine Learning (ML) modelling offers potential insights into water quality index prediction. This study employed advanced ML models to get potential insights into the prediction of water quality index for the Aak-Stream, an industrially polluted natural water resource in Pakistan with 19 input water quality variables aligning them with surrounding land use and anthropogenic activities. Six machine learning algorithms, i.e. Adaptive Boosting (AdaBoost), K-Nearest Neighbors (K-NN), Gradient Boosting (GB), Random Forests (RF), Support Vector Regression (SVR), and Bayesian Regression (BR) were employed as benchmark models to predict the Water Quality Index (WQI) values of the polluted stream to achieve our objectives. For model calibration, 80% of the dataset was reserved for training, while 20% was set aside for testing. In our comparative analyses of predictive models for water quality index, the Gradient Boost (GB) model stood out the fittest for its precision, utilizing a combination of just seven parameters (chemical oxygen demand, total organic carbon, oil & grease, Ammonia- nitrogen, arsenic, nickel and zinc), surpassing other models by achieving better results in both training ( $R^2 = 0.88$ , RMSE = 7.24) and testing ( $R^2 = 0.85$ , RMSE = 8.67). Analyzing feature importance showed that all the selected variables, except for  $\text{NO}_3^-$ , N, TDS and temperature had an impact on the accuracy of the models predictions. It is concluded that the application of machine learning to assess water quality in polluted environments enhances accuracy and facilitates real-time tracking, enabling proactive risk mitigations.



## **Assessing Soil Pollution around Aik-Stream: Utilizing Advanced Methods for Ecological Risk Assessment**

### **3.1 Introduction**

Soil pollution caused by heavy metals is a pressing issue in areas with industrial activities and urban growth (Zwolak, Sarzyńska, Szpyrka, & Stawarczyk, 2019). Because processes cannot break them down, heavy metals remain in the environment for some time, even after release (Ejaz et al., 2023a; Y. Huang et al., 2019). People can come into contact with these substances through avenues such as touching contaminated soil, directly consuming food and drinks that have been produced on the contaminated land, and inhaling dust that has been tainted (Rajesh KUMAR Mishra, Mohammad, & Roychoudhury, 2016). Long-term exposure to these pollutants can have consequences on our health, leading to conditions such as cancer, organ damage, and neurological disorders (Rodrigues & Römkens, 2018). Soil pollution not only affects humans but also has an impact on plants and animals. Plants absorb these pollutants through their roots, which leads to growth abnormalities and lower yields, which ultimately affects the quality of the produced food (Ahmad, Khan, Ali, Fatima, & Ali, 2019; Domene, Ramírez, Solà, Alcañiz, & Andrés, 2009). Moreover, this heavy metal pollution has the potential to spread through the food chain, affecting animals that rely on these plants for sustenance (Notten, Oosthoek, Rozema, & Aerts, 2005). The Global Soil Pollution Report for 2023 evaluates the presence of metal contamination in regions, emphasizing the importance and severity of this issue (Grafkina, Pitryuk, & Goryacheva, 2023). This particular issue highlights a dilemma that frequently arises when the growth of industries conflicts with the need to preserve the environment (Gautam et al., 2023). Over the years, researchers have increasingly focused on understanding the factors contributing to soil contamination. Regular monitoring is essential for preserving soil quality. It impacts ecosystem conservation, human health promotion, pollution management, and policy development (Ciobanu, 2013; Z. Li et al., 2019). However, traditional monitoring techniques alone are often inadequate. It is now widely acknowledged that incorporating advanced models or frameworks is crucial to assessing potential environmental hazards (S. Yang et al., 2023). Furthermore, soil-dwelling organisms vital for nutrient cycling and maintaining soil structure are also at risk (Van Straalen, 1996). Another significant concern is



the contamination of groundwater sources (Gillispie, Sowers, Duckworth, & Polizzotto, 2015). Pollutants from the soil can leach into these water sources, spreading the contamination over wide areas and endangering water quality for large populations (Arias-Estévez et al., 2008). Additionally, soil pollution adversely affects soil fertility, decreasing agricultural productivity (Saha et al., 2017). This has economic implications and impacts food security, particularly in agrarian societies (Elbana, Gaber, & Kishk, 2019). Soil pollution disrupts vital biogeochemical cycles like carbon and nitrogen, affecting the overall environmental balance and leading to broader ecological issues (Rajesh KUMAR Mishra et al., 2016). The deterioration of ecosystem services, such as nutrient cycling, water filtration, and habitat provision, is another crucial aspect of soil contamination (Cachada, Rocha-Santos, & Duarte, 2018).

### ***3.1.1 Soil Pollution by Industrial Wastewater: A Global Issue***

The issue of soil degradation due to industrial effluent is a primary worldwide environmental concern. Industrial wastewater often contains heavy metals and other hazardous materials that can seep into the soil and cause significant changes to the composition and health of the soil (Alnaimy, Shahin, Vranayova, Zelenakova, & Abdel-Hamed, 2021). This issue is more severe in underdeveloped and developing countries since insufficient waste treatment facilities exacerbate the situation. Untreated industrial effluents are frequently used in agriculture in these areas, which causes the soil to become contaminated with toxic substances (Zwolak et al., 2019). These harmful compounds, therefore, have the potential to make their way up the food chain, endangering human health and seriously harming the ecosystem (U. Banerjee & Gupta, 2017). A study by (Cachada et al., 2018) highlights the seriousness of this problem, pointing out the long-term effects of soil pollution on ecosystems and human health.

Similarly, (Kadi, 2009) highlights the extensive application of untreated industrial wastewater in agriculture and its consequent effects. The Global Soil Pollution Report (2023) provides a comprehensive overview of the extent of this problem, noting that millions of sites worldwide are contaminated (Grafkina et al., 2023). This instance represents a broader issue where industrial development and environmental preservation frequently conflict, particularly in less developed areas (Gautam et al., 2023). The absence of strict laws and enforcement mechanisms in these areas allows for continued pollution, often with irreversible consequences (D. K. Frankel, 1994). The infiltration of hazardous substances seeping into the



soil not only degrades the quality of the soil itself but also impacts water resources, air quality, and biodiversity (Ariffin, 2019).

### ***3.1.2 Heavy Metals in Soil: An Escalating Environmental Concern***

Soil pollution with heavy metals is one severe environmental problem caused mainly by modern industrial activities and urbanization (Zwolak et al., 2019). Because these metals are not biodegradable, they remain in the environment after being discharged frequently as byproducts of many industrial operations (Y. Huang et al., 2019). Their presence in soil is of significant concern due to their high toxicity levels. The accumulation of heavy metals in the soil can cause significant reductions in soil microbial activity (Miskowiec, Laptas, & Zieba, 2015). This reduction, as documented in research by (Miskowiec et al., 2015), directly affects soil fertility, hindering plant growth and disrupting natural soil processes. The human health implications of soil heavy metal contamination are considerable and diverse. These metals can enter the human body in several ways, including direct skin contact with contaminated soil and ingesting contaminated food and drink (Zwolak et al., 2019). Once inside the body, heavy metals can lead to several chronic health issues, including kidney damage, neurological abnormalities, and an increased chance of developing different kinds of cancer (Wimalawansa, 2016). Research by (X. Li, Jiao, Xiao, Chen, & Chang, 2015) clarified how urgently strong pollution control policies and efficient soil remediation plans were needed to lessen the effects of these pollutants. The study also stressed the significance of environmental regulations targeting industrial emissions and waste disposal methods to prevent soil contamination. Research on heavy metal contamination in soil has been extensive, with numerous studies highlighting the sources, effects, and mitigation techniques for various heavy metals (Song et al., 2017). Here, we delve into detailed insights into crucial heavy metals that cause soil pollution.

#### ***3.1.2.1 Chromium (Cr)***

Chromium, especially in its hexavalent form, is primarily released from industrial activities like leather tanning, metal plating, and dye manufacturing. A study by (Dotaniya, Thakur, Meena, Jajoria, & Rathor, 2014) confirms these industries' significant role in poisoning chromium pollution in soil. Many studies have been conducted on the effects of chromium, especially on soil ecosystems (Aparicio et al., 2019). Chromium exposure leads to oxidative stress in plants, inhibiting growth and causing cellular damage. This effect affects



agricultural productivity and natural ecosystems (Ertani, Mietto, Borin, & Nardi, 2017). According to (Ranieri & Gikas, 2014), crops cultivated in contaminated soils may accumulate chromium, which could harm human health if ingested. Constant monitoring of soil chromium levels is essential for controlling soil contamination. (Rani, Arya, & Dwivedi, 2020) investigated the effects of various soil types on chromium mobility and bioavailability, offering insights into the disparities in environmental risk and toxicity in various geographic locations.

#### *3.1.2.2 Cadmium (Cd)*

Cadmium (Cd) is commonly released into the environment through industrial operations such as metal mining, refining, battery manufacturing, and some plastics (M Hutton, 1983). (M. T. Hayat, Nauman, Nazir, Ali, & Bangash, 2019) found that cadmium negatively impacts soil microbiota, which reduces microbial diversity and impairs soil functioning. Cadmium has a significant impact on plant life. Exposure to cadmium causes oxidative stress in plants, inhibiting growth and leading to cellular damage, as shown in the study by (Hasan, Fariduddin, Ali, Hayat, & Ahmad, 2009). This affects natural ecosystems and has profound implications for agricultural productivity. There are serious health concerns for humans from cadmium in soils. Research has shown that crops cultivated in contaminated soils may accumulate cadmium, which could directly endanger human health when these products are consumed (HOCAOĞLU-ÖZYİĞİT & GENÇ, 2020). This demonstrates a direct connection between problems with food safety and soil contaminated with cadmium. Another study investigated the mobility and bioavailability of cadmium in soil, providing insight into the differing levels of environmental risk and toxicity in various geographic locations (McLaughlin & Singh, 1999).

#### *3.1.2.3 Copper (Cu)*

Copper (Cu), an extensively used metal in numerous industries and agricultural techniques, has become an environmental concern due to its potential to contaminate soils. The mining and smelting processes, the electrical and electronic sectors, and the application of chemicals are the main causes of pollution (Raoof Mahmood, Alheety, Asker, Zyaad Tareq, & Karadağ, 2019). A study by (Simonova & Cheglacova, 2017) revealed that elevated copper levels can cause soil acidification and reduce vital nutrients, upsetting the soil's nutrient balance. The impact of this mismatch on plant development and soil fertility may be



extensive. Excessive copper concentrations can be toxic to soil bacteria and fungi as these microorganisms are crucial to soil health, aiding in organic matter decomposition and nutrient cycling. Their decline can, therefore, harm soil ecosystem functions (Klimek & Niklińska, 2007). Excessive uptake of copper can lead to phytotoxicity, as evidenced in the research by (Lamb, Naidu, Ming, & Megharaj, 2012), where high copper levels were linked to reduced plant growth and poor crop yields. This issue affects not only plant health but also agricultural sustainability. Regarding human health, the potential for copper to enter the food chain through crop absorption is a severe issue. A comprehensive study by (Karim, 2018) indicated that long-term exposure to copper-contaminated foods could lead to health problems like liver and kidney damage and gastrointestinal disturbances. The mobility of copper in soils and its availability to plants depends mainly on soil pH, organic matter content, and texture, as detailed in a study by (Laurent et al., 2020). These factors determine how much copper is retained in the soil versus how much is available for plant uptake, thus influencing environmental and health risks.

#### 3.1.2.4 Arsenic (As)

Arsenic (As) primarily originates from natural sources and human activities such as industrial processes and arsenic-containing pesticides and fertilizers (Garelick, Jones, Dybowska, & Valsami-Jones, 2008). Different forms of arsenic in soil have varying levels of toxicity. A study by (Morin & Calas, 2006) showed that inorganic forms of arsenic are more toxic than organic forms. The persistence of arsenic in soils, even decades after the cessation of arsenic inputs, was the focus of a study (Shrivastava, Ghosh, Dash, & Bose, 2015). They discovered that arsenic could remain in the soil for extended periods, impacting plant health and soil quality even after the initial source of contamination had been eliminated. Arsenic contamination also affects soil fauna, including insects, earthworms, and soil microbes (Moriarty, Koch, Gordon, & Reimer, 2009). The interactive effects of arsenic with other soil contaminants have been an emerging area of research. According to a study by (Tu & Ma, 2003), the toxicity and mobility of arsenic in soil can be altered by the presence of other heavy metals, such as lead and cadmium, creating complex risk situations. Recent research has investigated how soil arsenic behavior is affected by climate change. According to research by (Bondu, Cloutier, Rosa, & Benzaazoua, 2016), climate change, such as higher temperatures and rainfall, may increase arsenic's mobility in soils and raise the possibility of groundwater contamination.



### *3.1.2.5 Lead (Pb)*

Lead (Pb) soil pollution, a significant environmental issue, arises from various sources, including industrial activities, burning leaded gasoline, and deterioration of lead-based products (Havugimana et al., 2017). According to research by (Abbaszade et al., 2022), lead pollution in soil can significantly change its chemical characteristics, including pH and nutrient availability. Ecosystem health is impacted by this change, which has a detrimental effect on plant development and lowers soil fertility. The toxicity of lead has a significant impact on soil microorganisms, which are essential to the health of the soil. According to a study by (Oyewole et al., 2019), high lead concentrations in soil can inhibit the diversity and activity of soil fungi and bacteria, lowering the soil quality and its ability to support plant life. Lead contamination of soil is harmful to plant development and reduces agricultural yield. Lead exposure has been linked to root damage, inhibition of photosynthesis, and decreased plant growth, all of which impact crop production and food quality (Zulfiqar et al., 2019). A study emphasized the dangers of eating crops cultivated in lead-contaminated soils, including the possibility that (Pb) might build up in the body and result in several health problems, including cardiovascular disease and cognitive decline (Alengebawy, Abdelkhalek, Qureshi, & Wang, 2021).

### *3.1.2.6 Mercury (Hg)*

Mercury (Hg) soil pollution is a severe environmental problem that significantly affects ecosystems and human health. Mercury can build up in soil and cause a variety of environmental problems. It is mainly emitted by mining, coal combustion, and industrial processes (Teng et al., 2020). A study revealed that even minute concentrations of mercury in soil can cause impairments to photosynthesis, water absorption, and nutrient absorption in plants, resulting in stunted growth and decreased agricultural yields (Gworek, Dmuchowski, & Baczewska-Dąbrowska, 2020). A growing body of research has focused on the relationship between mercury and other soil contaminants. According to a study (Alekseev & Abakumov, 2021), the mobility of mercury in soil can be increased by organic pollutants like polycyclic aromatic hydrocarbons (PAHs), which increases the possibility that the metal will be absorbed by plants and leak into groundwater. Understanding the role of soil microbiota in mercury methylation has been a significant focus. (W.-L. Tang et al., 2020) found that specific soil microbial communities could methylate mercury, highlighting the significance of microbial ecology in reducing mercury hazards. Novel approaches to mercury cleanup in soil



are being investigated using carbon-rich biochar to encapsulate mercury in soil and lessen its bioavailability and environmental impact (W. Chen et al., 2023). Long-range air transport of mercury was traced by (Yevugah, Darko, & Bak, 2021), demonstrating how industrial emissions in one place can cause mercury deposition in distant, even remote, areas.

#### *3.1.2.7 Nickel (Ni)*

Nickel (Ni) soil pollution, while less frequently discussed than other heavy metals, has been the subject of important research due to its increasing presence in the environment, primarily from industrial operations and waste disposal (El-Naggar et al., 2021). Nickel can build up in soils, particularly in locations close to industrial sites (A. Kumar et al., 2021; X.-Y. Zhou & Wang, 2019). According to a study by (X.-Y. Zhou & Wang, 2019), nickel concentrations in industrial soils can be several times greater than background levels and can last for a long time because of nickel's poor mobility. Nickel pollution can affect the composition of microbial communities and decrease microbial diversity, which can influence soil health and nutrient cycling (Minari et al., 2020). Nickel exposure in soil can lead to reduced plant growth, chlorosis (yellowing of leaves), and decreased photosynthetic efficiency. This may significantly impact agriculture production in contaminated areas (Hassan et al., 2019). Acidic soil conditions can increase the mobility and bioavailability of nickel, offering more significant dangers to plants and groundwater (El-Naggar et al., 2021). Exposure to soil contaminated with nickel provides health hazards; persons who are exposed to high quantities of nickel over an extended period may experience skin irritation, lung problems, and other health concerns (Genchi, Carocci, Lauria, Sinicropi, & Catalano, 2020).

#### *3.1.2.8 Zinc (Zn)*

Zinc (Zn), a necessary element micronutrient for plants and microorganisms, can have several harmful effects on soil and the broader ecosystem when present in excessive concentrations (Cao, Xie, & Hou, 2022). Soil degradation can result from high zinc levels that drastically alter the soil's chemical features, including pH and nutrient balance (Gautam et al., 2023). High zinc levels suppress microbial development and activity, which lowers soil fertility and interferes with ecological processes (B. Tang, Xu, Song, Ge, & Yue, 2022). In excessive zinc conditions, plant nutrient intake can be stuck, leaves may become yellow due to chlorosis, and photosynthesis may be decreased, resulting in stunted growth and lower crop yields (Natasha et al., 2022). Zinc-contaminated soils can cause plants to absorb the



metal, which can then enter the food chain and endanger the health of both humans and animals (Zwolak et al., 2019). High zinc concentrations can also impact soil fauna, including insects and earthworms. Zinc contamination can lower these creatures' survival and reproduction rates, affecting soil structure and health (chatelain, 2023). Each of these metals poses unique challenges and hazards to soil health and the broader environment, emphasizing the need for continued research and implementation of effective mitigation measures.

Artificial intelligence (AI) has recently been perceived as a notable advancement, particularly in the advancement of algorithms used to assess soil quality through intricate data analysis (Gautam et al., 2023). One of the AI tools available is the Self Organizing Map (SOM), an algorithm that draws inspiration from the networks found in the human brain. It is well known for its capability to group and display data in dimensions (Guagliardi, Astel, & Cicchella, 2022; Xiang et al., 2022). Research has indicated that SOMs can overcome the limitations encountered by classification techniques, providing a comprehensive and nuanced perspective in analyzing data (Olawoyin, Nieto, Grayson, Hardisty, & Oyewole, 2013). Self-organizing maps (SOMs) function by finding similarities in input vectors and grouping them in an environment that does not require supervision (Licen, Astel, & Tsakovski, 2023; Olawoyin et al., 2013). In contrast, approaches that depend on statistics and human intuition, often classified by particular rules or formulas, are vulnerable to human prejudice and cumulative mistakes. SOMs present a self-governing, more impartial methodology (Mele & Crowley, 2008). By enabling computers to analyze and classify data, self-organizing maps (SOMs) decrease the influence of subjectivity and statistical inaccuracies (Guagliardi et al., 2022; Xiang et al., 2022). The central objective is to evaluate the effectiveness of self-organized maps (SOMs) in classifying and understanding soil pollution caused by metals. Furthermore, using multivariate analysis and integrative modeling, such as Structural Equation Modeling (SEM), represents an advancement in ecological research (Fan et al., 2016). This study provides insights into understanding interactions. It presents these methodologies as tools for deciphering ecological relationships and identifying and linking environmental risk factors, such as pollution sources (Ilacqua et al., 2007). Comprehensively analyzing data is vital for developing management and conservation strategies. These approaches represent a change in how we interpret and use data (Malaeb, Summers, & Pugeseek, 2000; A. u. Rahman et al., 2021).



This study addresses soil pollution in an environmentally vulnerable area near the polluted Aik Stream in Sialkot, Pakistan. This area has been dramatically affected by disturbances and has suffered harm due to pollution caused by toxic metals such as lead, cadmium, nickel, and mercury (A. Qadir & Malik, 2009). Over the three decades, the environmental health of this region has been increasingly threatened by the expansion of industrialization and urban development. The pollution in this region has negatively impacted the soil quality, posing a significant risk to its overall health (Malik, Husain, & Nazir, 2010). In the region, we can observe an absence of comprehensive monitoring approaches that accurately assess the environmental hazards posed by heavy metals. To effectively address the contamination issue and its impact on the environment, it is essential to understand its seriousness and potential ecological repercussions. This understanding will enable us to develop strategies for managing and remedying the situation. To address this issue, the study establishes its objectives to 1) Evaluate and quantify the intensity of heavy metals in soil by different pollution indices 2) Categorized the study area based on varying levels of soil pollution with the help of a Self-Organizing Map (SOM) technique. 3) Analyze the potential ecological risks of heavy metals in the soils and identify the primary sources behind these risks using Structure Equation Modeling (SEM). The main goal of this study is to gain insight into the impacts caused by contamination and identify the primary sources responsible for heavy metal pollution in various parts of the study region. In investigating soil pollution, this study aligns with Sustainable Development Goal 15, which focuses on achieving 75% healthy soils by 2030 (Nations, 2015). It highlights the need to tackle soil contamination for land utilization and the preservation of ecosystems.

## **3.2 Materials and Methods**

### ***3.2.1 Study Design***

Soil samples were collected from 150 locations across 30 stations along the stream. At each spot, three samples were combined to form a soil sample weighing 1.5 kg. These smaller samples were collected from areas within the grid region ranging in depth from 3 to 15 cm. The soil was collected with a hand trowel, stored in clearly labeled polyethylene bags, and then transported to the laboratory for detailed physical and chemical analysis.



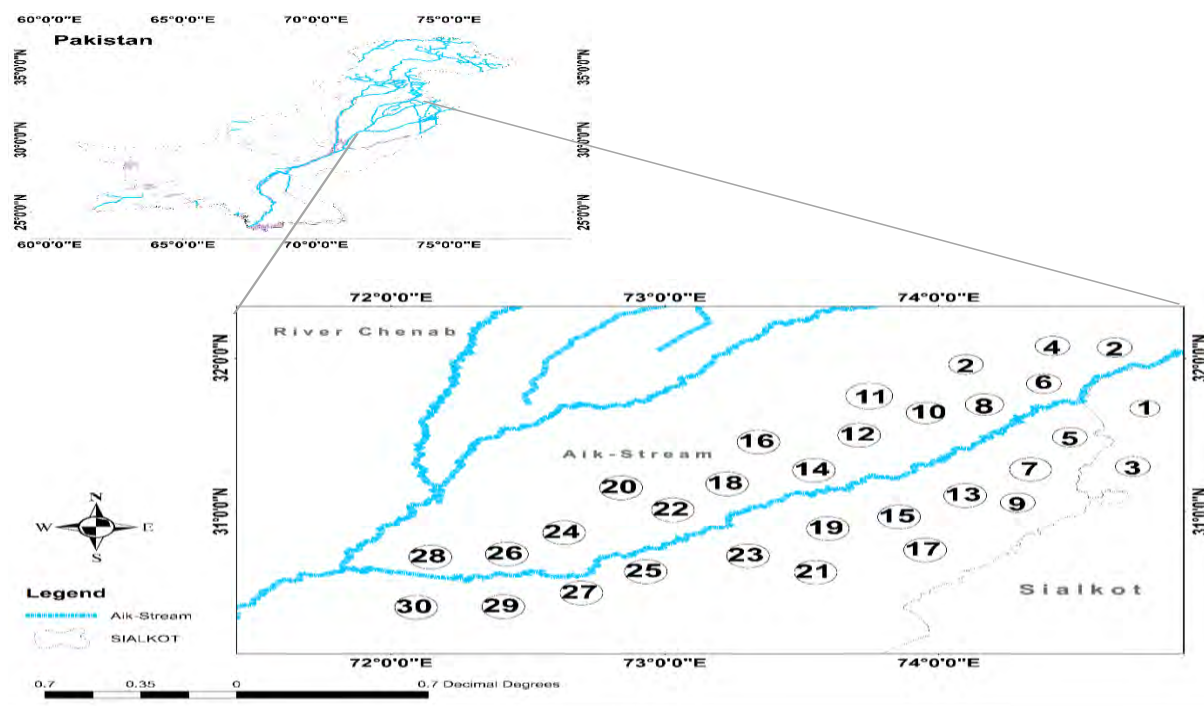


Figure 3. 1 Map Illustrating the Industrial City of Sialkot, Pakistan, Alongside the Aik Stream and Soil Sampling Stations (Source: Arc map 10.5)

### 3.2.2 Analytical Procedures in the Laboratory

The samples collected were transported to the Plant Ecology Laboratory at Quaid-i-Azam University, Islamabad, where they were subjected to air-drying for approximately two weeks. The samples were prepared for analysis after grinding and sieving through a 2 mm mesh to remove larger debris. The analysis aimed to assess various critical physicochemical and biochemical characteristics. In the laboratory, a range of water quality parameters was determined, including pH, temperature ( $^{\circ}\text{C}$ ), electrical conductivity (EC  $\mu\text{S}/\text{cm}$ ), total dissolved solids (TDS  $\text{mg}/\text{L}$ ), organic matter (OM %), phosphorus (P  $\text{mg}/\text{kg}$ ), and potassium (K  $\text{mg}/\text{kg}$ ). Additionally, the concentrations of several metals (Zn, Ni, Cu, Cr, Pb, As, and Hg) were determined, each expressed in  $\text{mg}/\text{kg}$ .

TDS, EC, and pH were measured using TDS, EC meters, and a pH meter (Russel RL060P). A soil suspension, prepared in a 1:9 ratio of soil to distilled water, was stirred for 60 seconds at 10-minute intervals over 30 minutes before the pH, EC, and TDS measurements were taken. The Loss on Ignition (LOI) method determined organic matter. The soil sample was dried at  $105^{\circ}\text{C}$  until a constant weight was reached. Subsequent incineration in a muffle furnace at  $360^{\circ}\text{C}$  for two hours led to the combustion of organic matter. The residue, upon cooling, was weighed, and the organic matter content was calculated based on the weight loss.



Phosphorus levels in the soil were determined using either the Bray-1 method for soils with a pH less than 7.0 or the Olsen method for soils above 7.0. Both methods involved the soil being mixed with the respective reagent, the supernatant being filtered, and the phosphorus concentration is measured using a spectrophotometer based on the colour intensity of the solution. The potassium levels in the soil were evaluated by extracting potassium using an ammonium acetate solution, followed by the filtration of the resulting solution. The potassium concentration was then measured using flame photometry. Heavy metals (Zn, Cu, Ni, Cr, As, Cd, Pb, and Hg) were detected using atomic absorption spectrophotometers, specifically the VARIAN AA240FS model (Cantle, 1986). For the analysis of metals, the soil samples underwent acid digestion. This digestion process was conducted using nitric acid (69% purity), perchloric acid (65% purity), and hydrochloric acid (28% purity). A measure of 0.5 g of each soil sample was digested in a solution containing 10 ml of concentrated nitric acid, 3 ml of perchloric acid, and 2 ml of hydrochloric acid. This mixture was subjected to a 30-minute digestion process using a microwave-accelerated reaction system (MARS, CEM). After the digestion, the soil samples were analyzed for metal content using a Fast Sequential Atomic Absorption Spectrometer (Varian FS-AA-240). This spectrometer allowed for precisely determining various metal concentrations in the digested soil samples.

### **3.2.3 Statistical analysis**

Statistics analyses were performed using Microsoft Excel 2013, R software (version 4.0.2), and Python (version 3.5).

#### **3.2.3.1 The selected indices**

In the study area, soil contamination is assessed using two pivotal indices: the Contamination Factor (CF) and the Degree of Contamination (DC). The Contamination Factor (CF) is calculated using the formula (Eq.3.1) (Ravankhah, Mirzaei, & Masoum, 2015).  $C_n$  represents the concentration of a specific heavy metal at a sampling site, and  $C_b$  is the permissible level for that particular heavy metal. This index serves as a focused metric, with CF values below 1 indicating non-pollution, values between 1 and 3 signifying moderate pollution, and values exceeding 3 indicating considerable to highly polluted sites within the study area (Amadi & Nwankwoala, 2013).

$$\text{Contamination Factor } CF = \frac{C_n}{C_b} \dots \dots \dots (3.1)$$



Additionally, the Degree of Contamination (DC) is determined by summing up the individual Contamination Factors (CF<sub>i</sub>) for all the heavy metals under investigation. The formula for DC is shown in (Eq. 3.2). DC provides a comprehensive overview, offering an aggregated assessment of contamination levels across multiple metals (El-Amier, Elnaggar, & El-Alfy, 2017). A DC less than 6 suggests low contamination, a range of 6 to 12 implies moderate contamination, 12 to 24 indicates considerable contamination and a DC exceeding 24 signifies a high level of contamination (Q. Jiang, He, Ye, & Christakos, 2018).

$$\text{Degree of Contamination } DC = \sum CF_i \dots \dots \dots (3.2)$$

### 3.2.3.2 Potential Ecological Risk Index (PERI)

The Potential Ecological Risk Index (PERI) method, pioneered by Swedish scientist Lars Hakanson in 1980 (Hakanson, 1980b), is designed to assess the extent of pollution and potential ecological damage caused by heavy metals in soil and sediments. This approach is rooted in sedimentology principles, considering heavy metals' characteristics and environmental behavior. The PERI method incorporates a toxicity response coefficient, establishing a link between ecological impacts, environmental consequences, and the toxicology of heavy metals. Consequently, the assessment focuses on comprehending heavy metal toxicity's overall migration and transformation patterns in soil and sediment (Ke et al., 2017). It also considers the sensitivity of the evaluation area to heavy metal pollution and variations in regional background values of heavy metals. Additionally, PERI factors in characteristics such as biological availability, relative contribution, and geographical spatial differences. As a comprehensive indicator, PERI provides valuable insights into the potential ecological impact of heavy metals (Suresh, Ramasamy, Meenakshisundaram, Venkatachalapathy, & Ponnusamy, 2011). This information is a basis for environmental improvement initiatives and offers a scientific reference for promoting a healthier living environment. The calculation formula for PERI is detailed in (Eq.3.3 and 3.4).

$$PER = TRE.CF \dots \dots \dots (3.3)$$

$$PERI = \sum PER \dots \dots \dots (3.4)$$

In equations 3.4 and 3.5, CF represents the pollution coefficient attributed to heavy metals, denoting the measured concentration of heavy metals in milligrams per kilogram (mg/kg). This study establishes permissible levels using WHO values for heavy metal elements in the soil of Aik Stream (Organization, 1996). PERI, identified as the potential ecological risk



coefficient for heavy metals, involves the toxicity response coefficient of the pollutant, denoted as TRE. TRE primarily reflects the toxicity level of heavy metals and the environment's sensitivity to heavy metal pollution (Ke et al., 2017). Consequently, PERI is the potential ecological hazard index for various heavy metals within a specific sampling point area (Tian, Huang, Xing, & Hu, 2017).

#### 2.2.4 Heavy Metal Absorption from Water to Soil

In this study we used soil absorption coefficient (Kd) the soil water partition coefficient to determine the hypothesized model the extent of water and soil pollution within ecosystems influenced by industrial pollution. Data used in this study is collected from different polluted zones from water and soil. The following equation used to determine the soil-water partition coefficient (Kd) for heavy metals:

$$Kd = \frac{\text{heavy metal concentration in soil}}{\text{heavy metal concentration in water}}$$

The soil-water partition coefficient, abbreviated Kd, is measured in mg/kg/mg/L (milligrams per kilogram/milligrams per liter). The amount of the heavy metal present in the soil, which is commonly expressed in milligrams per kilogram (mg/kg). The amount of the heavy metal present in the water, usually expressed in milligrams per liter (mg/L).

Table 3.1 The different types of models, the rank of the values, and the degree of soil contamination.

Index	Description	Index value	Contamination degree	source
<b>Contamination Factor (CF)</b>	The Contamination factor is the ratio obtained by dividing the concentration of each metal in the soil by the permissible value.	CF < 1	Low degree of Contamination	(Ravankhah et al., 2015)
		1 < CF < 3	Moderate degree of Contamination	
		3 < CF < 6	Considerable degree of Contamination	
		CF > 6	Very high degree of contamination	
<b>The Degree of Contamination (DC)</b>	The contamination level may be classified based on their intensities on a scale ranging from 1 to 6 (0 none, 1 none to	DC < 6	Low	(Yeh et al., 2020)
		6 < DC < 12	Moderate	
		12 < DC < 24	Considerable	



	medium, 2 polluted).	moderate, 4	DC > 24	Very high
<b>Potential Ecological Risk</b>	Potential ecological risk is obtained by the multiplication of CF with a toxicity of the specific heavy metal.	PER < 1	Low	(Hakanson, 1980b)
		2 < PER < 5	Moderate	
		5 PER < 20	Considerable	
		PER > 20	High	
<b>Potential Ecological Risk Index</b>	the PERI provides a fast and simple quantitative value for environmental assessment	PERI < 150	Low	(Hakanson, 1980b)
		150 < PERI < 300	high	
		PERI > 600	Very high	

### 3.2.5 Self-Organizing Map (SOM)

The self-organizing map (SOM) is an unsupervised learning artificial neural network model that categorizes input patterns by determining the optimal reference vector set (Kohonen, 1990). The structural framework for the SOM neural network learning is illustrated in Figure 3.2 (Ghaseminezhad & Karami, 2011). The SOM process unfolds as follows:

#### 3.2.5.1 SOM neural network learning step

The most conventional output composition for the self-organizing map (SOM) network is a two-dimensional planar organization, often described as a 2D map. Its layout aligns with the signal-processing pattern found in the cerebral cortex. Each neuron is interconnected laterally with its neighboring neurons within this output plane, forming a structured checkerboard pattern.

**Step 1:** variables and parameters.

The input vector is  $X(n) = (X_1(n), X_2(n), \dots, X_n(n))^T$

The weight of the vector is  $Wi(n) = (W_{i1}(n), W_{i2}(n), \dots, X_{in}(n))^T$

Set the number of iterations as N

**Step 2:** Data initialization

Initialize the weight vector  $Wi$  and set the initial learning rate  $\alpha_0$ ; normalize the initial value  $Wi(0)$  of the weight vector and all the input vectors  $X$ .

$$X' = \frac{X}{||X||} = \frac{(X_1, X_2, X_n)^T}{(X_1^2 + X_2^2 + X_3^2 \dots X_n^2)^{1/2}} \dots \dots \dots (3.6)$$



$$W(0)' = \frac{Wi(0)}{||Wi(0)||} \dots \dots \dots (3.7)$$

Where  $||Wi(0)|| = (\sum_{j=1}^n [Wij(0)]^2)^{1/2}$ ,  $||X|| = (\sum_{j=1}^n (xj)^2)^{1/2}$  they are the norms of the weight vector and the input vector, respectively.

**Step 3:** Select the training sample  $X'$  from the input space.

**Step 4:** Calculate the Euclidean distance  $d_i$  between  $i'$  and  $X'$ .

$$d_i = \sum_{j=1}^n [(Xi - Wij)^2]^{1/2} \dots \dots \dots (3.8)$$

$i = 1, 2, 3 \dots m$ , is the distance among the  $m$  Euclidean between the weight and training samples.

**Step 5:** Approximate matching

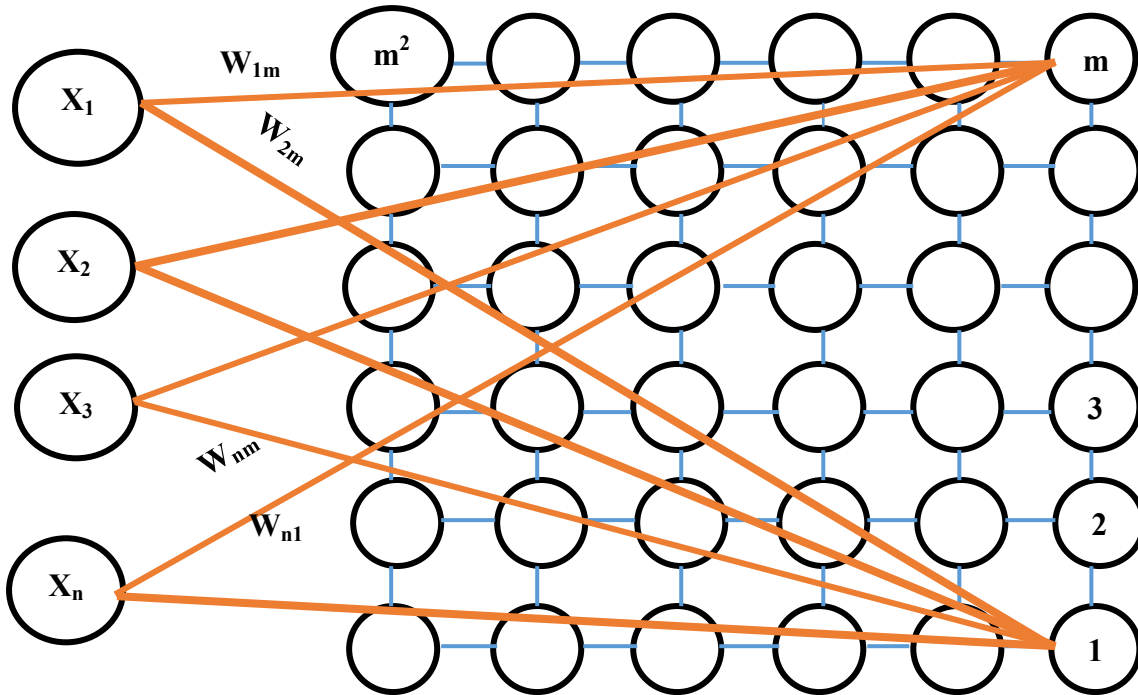


Figure 3. 2 Basic structure of the SOM network

The principle of minimum Euclidean distance is:

$$||X' - Wc'|| = \min_i ||X' - Wi'|| = \min_i [d_i], i = 1, 2, 3 \dots m \dots \dots \dots (3.9)$$

Win neuron  $c$  through competitive learning.



**Step 6:** Iteration.

For the excited neurons in the topological neighborhood  $N_i(x)(n)$  of the winning neurons, the weight vector of the neurons is iterated, resulting in a cooperative and iterative process for the neurons.

$$\begin{cases} W_j(n+1) = W_j(n) + \gamma(n)h_j \cdot i(x)(n)(X - W_j(n)) & j \in N_{(i)(n)} \\ W_j(n+1) = W_j(n) & j \notin N_{(i)(n)} \end{cases} \dots \dots \dots (3.10)$$

**Step 7:** Iterate the learning rate  $\eta$  and topology domain and renormalize the weights after learning.

$$\gamma_n = \gamma_0 \exp\left(-\frac{n}{\tau_2}\right), n = 0, 1, 2, 3, \dots, N \dots \dots \dots (3.11)$$

$$\sigma(n) = \sigma_0 \exp\left(-\frac{n}{\tau_2}\right), n = 0, 1, 2, 3, \dots, N \dots \dots \dots (3.12)$$

$$W'_{(n+1)} = \frac{Wi_{(n+1)}}{\|Wi(n+1)\|} \dots \dots \dots (3.13)$$

**Step 8:** Judge whether the number of iterations  $n$  exceeds  $N$ . If  $n \leq N$ , return to step 3 otherwise the iterative process ends.

*3.2.5.2 Determination of neuron size in the SOM competitive layer*

In this research, the neural network for training consists of 150 samples, with each network containing data from eight soil heavy metal elements. The study adopts (Vesanto et al., 2000) estimation formula to determine the scale of the self-organizing map. The initial scale size is estimated using Vesanto Formula  $5\sqrt{n}$  (where 'n' represents the number of samples). Subsequently, the optimal map scale size is determined by minimizing Quantitative Error (QE) and Structural Error (TE). Quantitative Error (QE) quantifies the average distance between each data vector and the best-matched neuron. Conversely, Structural Error (TE) assesses the map's ability to accurately preserve the data's topology. TE is calculated based on the proportion of samples not adjacent to the first and second Best Matching Units (BMU) relative to the total number of input samples.

*3.2.5.3 Determination of the best classification number*

Determining the optimal classification number in the Self-Organizing Map (SOM) network is critical in achieving meaningful spatial clustering. Unlike automatic methods, the SOM network requires enhancement by integrating clustering techniques. This study combines the K-means method and the Davies-Bouldin Index to identify the best



classification number. (Nakagawa, Yu, Berndtsson, & Hosono, 2020; Qu, Shi, Liang, Wang, & Han, 2021). The process begins with K-means clustering, where the number of categories varies from 2 to N, with 'N' representing the total number of input data samples. This step explores different clustering scenarios. Subsequently, the Davies-Bouldin Index is utilized to evaluate the quality of clustering outcomes. It calculates the average dissimilarity between clusters, aiding in determining the optimal classification number. Integrating these methods ensures that the SOM network achieves scientifically sound spatial clustering, enhancing its ability to identify the most suitable classification number for the given dataset (Günter & Bunke, 2003).

### 3.2.6 Methodology of Structural Equation Models (SEM)

In order to investigate the links between soil metals and various environmental factors such as wastewater, human activities, farming practices and weather conditions in the study area, we utilized linear structural equation modeling (LSEM) with the help of the 'lavaan' package. (Rosseel et al., 2017). In our approach, we conduct tests on the LSEM paths derived from our hypothetical model (Figure 3.3). Our model considered water quality, soil quality, anthropogenic pressure and climatic conditions as factors that can influence the presence of metals. In our analysis, we considered these factors to be influences. We used structural equation modeling (SEM) to integrate hypotheses and mechanisms into a model. Furthermore, the models were adjusted to include factors that capture errors, accounting for variable variations due to correlations.

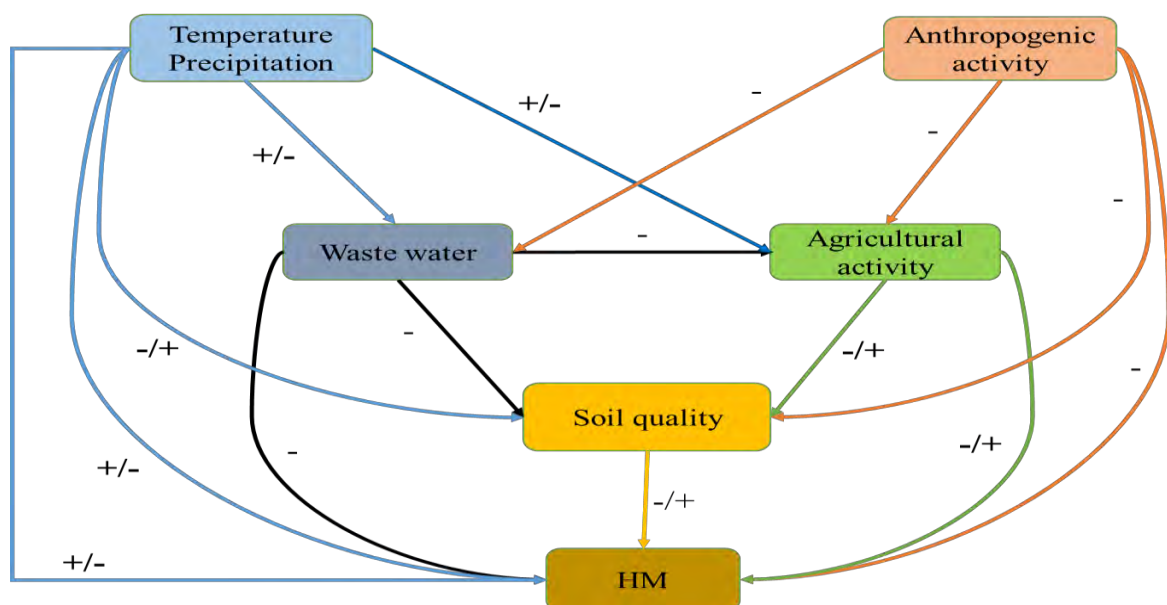


Figure 3.3 A conceptual model illustrating the associations between heavy metals and components in the Aik stream region, such as wastewater, anthropogenic pressure,



agricultural activities, climatic conditions, and soil qualities. The symbols (+ and -) show the directional influence of certain factors.

$$HMs = \alpha + \beta_0 SQ + \beta_1 WQ + \beta_2 AA + \beta_3 Temp + \beta_4 preci + \beta_4 AP \varepsilon_i \dots \dots \dots (3.14)$$

$$HMs = \alpha + \beta_0 SQ + \beta_1 WQ + \beta_2 AA + \beta_3 Temp + \beta_4 preci + \varepsilon_i \dots \dots \dots (3.15)$$

$$HMs = \alpha + \beta_0 SQ + \beta_1 WQ + \beta_2 AA + \beta_3 Temp + \varepsilon_i \dots \dots \dots (3.16)$$

$$HMs = \alpha + \beta_0 SQ + \beta_1 WQ + \beta_2 AA + \varepsilon_i \dots \dots \dots (3.17)$$

$$HMs = \alpha + \beta_0 SQ + \beta_1 WQ + \varepsilon_i \dots \dots \dots (3.18)$$

$$HMs = \alpha + \beta_0 SQ + \varepsilon_i \dots \dots \dots (3.19)$$

In Equations 14-19, 'HMs' represent the heavy metals, 'SQ' denotes soil quality, 'WQ' stands for water quality, 'AA' signifies agricultural activity, and 'AP' refers to anthropogenic pressure and temperature and precipitation in the study area. The symbol 'βs' is used for the coefficients of the explanatory variables, 'α' for the intercept, and 'ε<sub>i</sub>' for the error term in the models.

### 3.3 Results

#### 3.3.1 Evaluation of Soil Parameters

Soil samples from the study area were analyzed for various parameters, including pH, temperature (°C), electrical conductivity (EC, μs/cm), total dissolved solids (TDS, mg/l), organic matter (OM, %), phosphorus (P, mg/kg), and potassium (K, mg/kg) and the concentrations of several metals (mg/kg), namely zinc (Zn), lead (Pb), nickel (Ni), copper (Cu), cadmium (Cd), chromium (Cr), arsenic (As), and mercury (Hg) (Table 3.2). Considerable variations in pH were detected across the sampled sites, with the values ranging from 6.55 to 9.20. In terms of EC, a range from 1.10 to 9.20 μS/cm was recorded. The lowest EC readings were associated with soil samples collected upstream, whereas the highest readings were observed in midstream regions. A deficiency in organic matter was identified, with the content ranging between 0.050 and 0.990. Total Dissolved Solids (TDS) varied between 51 and 757 mg/kg, while Phosphorus (P) and Potassium (K) contents ranged from 2.50 to 11.6 mg/kg and 243 to 744 mg/kg, respectively. There was substantial variation in the number of heavy metals in the soil. Chromium (Cr) concentration oscillated from 20.8 to 266 mg/kg, identifying two distinct sub-groups. In the soil samples analyzed, 75.37% exceeded 100 mg/kg toxicity for Cr. The Cd, Pb, and Zn amounts were noted to vary between 0.021 and 4.87 mg/kg, 2.62 and 99.8 mg/kg, and 10.2 and 81.5 mg/kg, respectively. Additionally, 60.37% of the samples exceeded the toxicity extent for Zn (50 mg/kg). The highest recorded



level of Pb was 1.8 times higher than its toxic limit, and the Cu level was 5.8 times greater than the toxic limit of 36 mg/kg in soil. In contrast, Arsenic (As) contents ranged from 0.678 to 1.98 mg/kg, showing a lower concentration trend in urban and semi-urban regions. It was noted that soil samples from urban regions exhibited more significant levels of heavy metals than those from semi-urban regions, suggesting an influence of urban activities on soil contamination levels. The elevated concentrations of heavy metals exceeded permissible limits, suggesting likely environmental risks. Notably, chromium exhibits the highest variability, evidenced by a substantial standard deviation 67.8, underscoring its spatial heterogeneity (Table 3.2).

Table 3.2 Summary statistics of soil samples from the study area detailing the sample size [N = 150 (3x)].

Soil Parameters	Mean	St.dev	Minimum	Maximum	Skewness	Kurtosis	Permissible limit	Toxicity
Temperature (°C)	44.1	0.804	2.84	44.1	-12.5	155	26.6	-
Soil moisture	0.630	0.540	0.630	0.630	-12.5	155	-	-
TDS (mg/L)	315	152	51	757	0.646	0.314	500	-
Saturation (%)	34.1	5.40	20.0	58.0	0.953	0.916	-	-
pH	7.54	0.533	6.55	9.20	0.973	0.80	7	-
EC (μS/cm)	4.58	3.79	1.10	17.9	1.94	3.46	110	-
OM (%)	0.39	0.550	0.050	0.990	-0.03	-1.40	0.05	-
P (mg/kg)	7.06	1.71	2.50	11.6	-0.13	0.206	30	-
K (mg/kg)	421	78.7	243	744	1.32	0.196	300	-
Cu (mg/kg)	25.7	14.9	5.34	65.1	0.632	-0.55	36	5
Zn (mg/kg)	38.4	17.5	10.2	81.5	0.555	-0.64	50	1
Cr (mg/kg)	80.2	67.8	20.8	266	1.27	0.555	100	2
Pb (mg/kg)	49.8	37.9	2.62	99.8	0.0996	-1.85	85	5
Cd (mg/kg)	1.38	1.22	0.021	4.87	0.983	0.983	0.8	30
As (mg/kg)	0.748	0.678	0.678	1.98	0.598	-1.28	10	0.5
Ni (mg/kg)	20.5	19.8	0.342	67.4	0.464	-1.49	35	5
Hg (mg/kg)	0.179	0.192	0.0080	0.928	0.928	1.98	1	0.27

### 3.3.2 Heavy Metal Contamination in the Soil

Heavy metal contamination in the soil samples was analyzed through two key indices: the Contamination Factor (CF) and the Degree of Contamination (DC). Significant insights into the extent and distribution of heavy metal were revealed by applying these indices. The concentration factor (CF) values for different heavy metals in soil samples exhibited a specific order: Cadmium (Cd) > Chromium (Cr) > Zinc (Zn) > Copper (Cu) > Lead (Pb) > Nickel (Ni) > Arsenic (As) > Mercury (Hg). These values indicate that the soil is highly



enriched with cadmium, especially in samples S72-S106, and also showing chromium, particularly in samples S100-S115. Zinc also exhibited high contamination in samples 103, 104, 105, 106, and 111. Copper presented moderate to considerable contamination across most locations, but samples 113 and 114 were highly contaminated regions. Arsenic and Mercury levels were generally low, suggesting minimal contamination, whereas Lead and Nickel showed considerable contamination. The study also measured the degree of contamination (DC) values; sites S83-S109 are identified as the most contaminated, exhibiting the highest DC values among all studied areas. In contrast, sites S1-S48 displayed the lowest DC values, and sites S125-S150 showed moderate DC values, as illustrated in (Figure 3.4, Table 3.3, and Appendix 3).

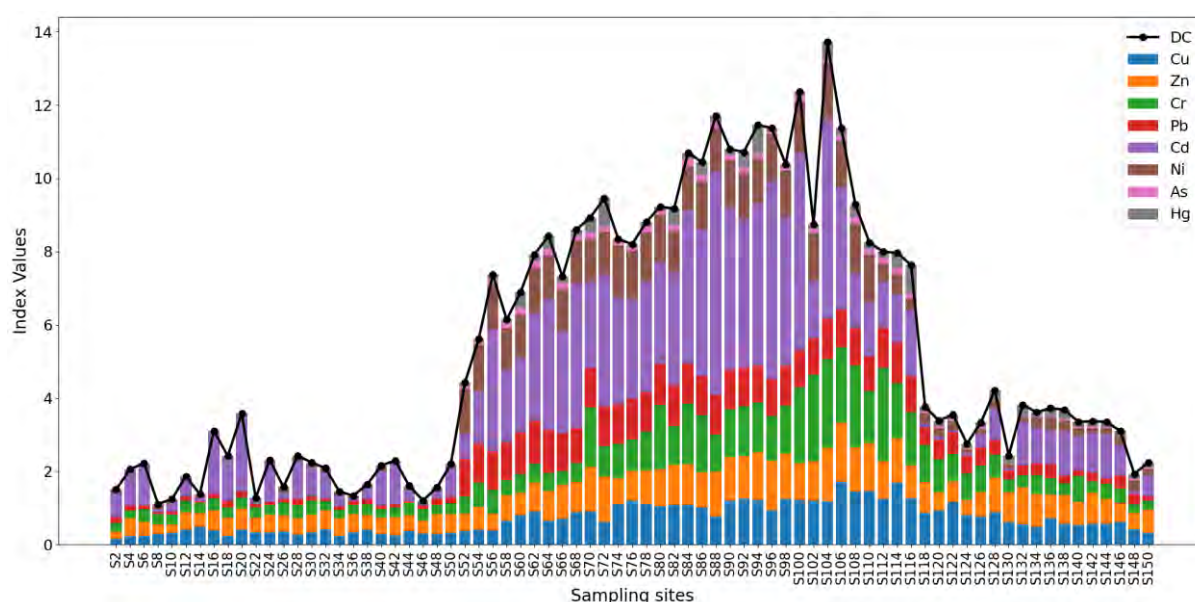


Figure 3. 4 Insights into the degree and distribution of heavy metal contamination in the vicinity of the Aik Stream area, focusing on two key indices: the Contamination Factor (CF) and the Degree of Contamination (DC).

Table 3.3 & Figure 3.5 displays the concentration of numerous heavy metals in soil over three distinct zones: the LPZ (Less Polluted Zone), MPZ (Moderate Polluted Zone), and HPZ (High Polluted Zone), with WHO-permitted levels. Specific parameters include heavy metals such as chromium (Cr), nickel (Ni), cadmium (Cd), lead (Pb), arsenic (As), mercury (Hg), and other elements such as copper (Cu) and zinc (Zn). Additionally, it measures the quantity of organic matter, electrical conductivity, total dissolved solids (TDS), phosphorus (P), potassium (K), pH (acidity/alkalinity), electrical conductivity, and pH. Table 3.3 and Fig. 3.5 represent the limits the World Health Organization (WHO) established for each metal, which serves as a global standard for evaluating substance safety and quality. These



measurements are crucial for environmental and health evaluations to identify pollution levels and potential dangers across different zones and eventually guide decisions and corrective measures. The results indicate that the concentration of heavy metals in the HPZ is greater than the WHO's pressable level.

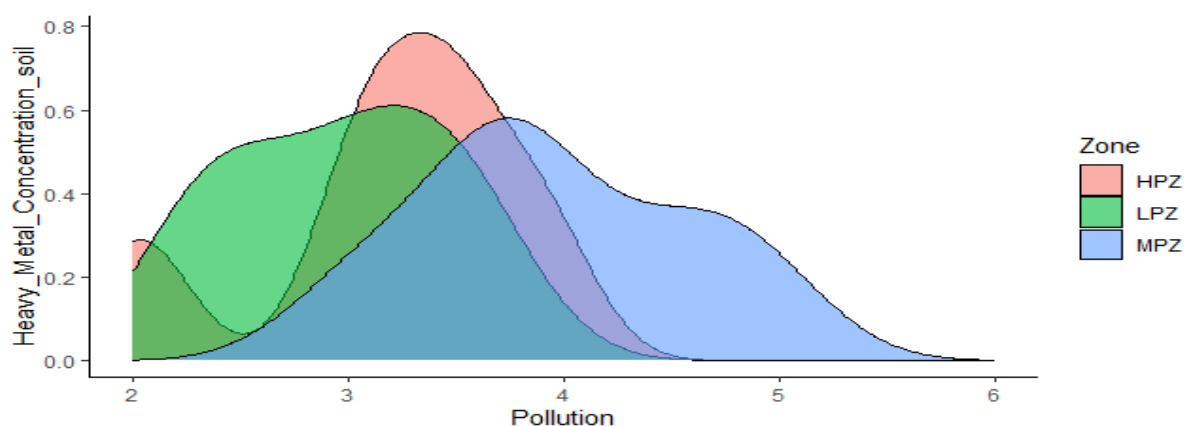


Figure 3. 5 Indicate the concentration of heavy metals in different zones (HPZ, LPZ, and MPZ) of the Aik stream.

The  $K_d$  values give essential information on how heavy metals behave in the environment. A higher  $K_d$  value indicates that the heavy metal has more potential to be absorbed by the soil and a lower propensity to migrate from water to soil. On the other hand, a lower  $K_d$  value denotes a heavier metal that is more mobile and prone to moving from water to soil. In order to estimate the possible environmental dangers linked to heavy metal contamination and to determine suitable management and remediation techniques, it is crucial to understand these transfer rates. Table 3.3 and Fig 3.6 shows the transfer rates of various heavy metals from water to soil in three different zones: Low Polluted Zone (LPZ), Moderate Polluted Zone (MPZ), and High Polluted Zone ( $K_d$  values) (HPZ). In the LPZ, the transfer rate is lower than in the other two zones. The higher transfer rate of Cr was 44.24, while the other metal had a lower absorption rate from water to soil in LPZ. However, Cr exhibits a substantially higher  $K_d$  value of 240.5894 in the HPZ, where the soil has a higher affinity for holding onto heavy metals, indicating mobility and increased adsorption onto the soil. Contrarily, cadmium (Cd) has lower  $K_d$  values in all zones, indicating a propensity for transfer from water to soil, particularly in the LPZ. Similarly, the transfer rates among all zones are also very low for copper (Cu) and zinc (Zn). It has a high  $K_d$  value of 97.29546 in the HPZ, showing a significant affinity for soil nickel (Ni). However, lower values in the MPZ and LPZ indicate reduced adsorption in these zones.



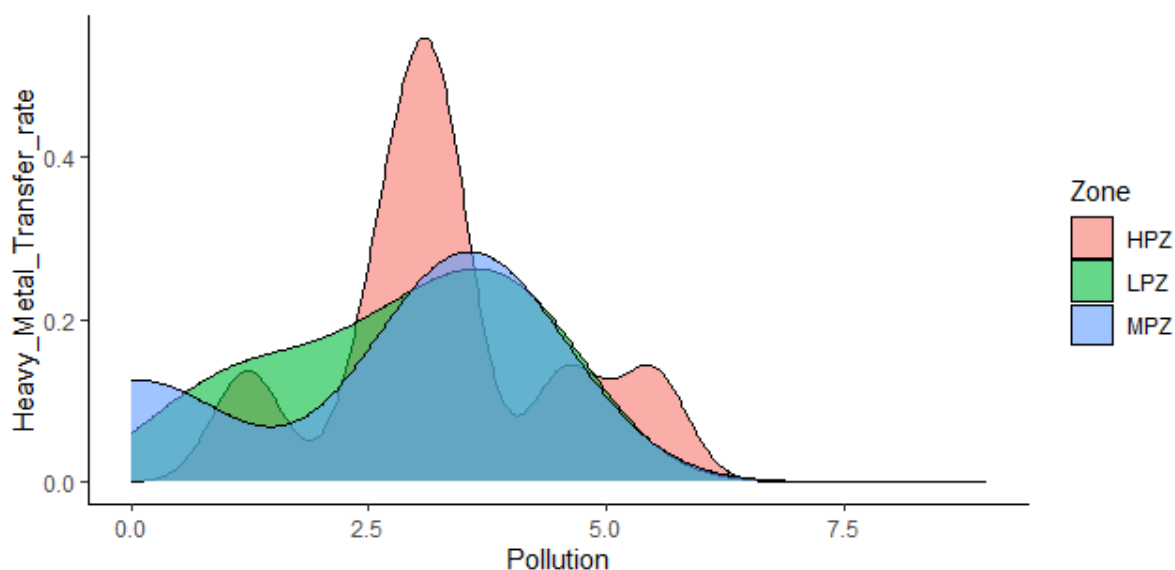


Figure 3. 6 Heavy metals transfer rate from water to soil in different zones of the Aik stream.

Table 3. 3 The transfer rate of heavy metal from water to soil

Parameter	Measured Valued (mg/kg)/(mg/L)		
	LPZ	MPZ	HPZ
Cr (Kd)	44.24825	73.03101	240.5894
Ni (Kd)	73.29546	22.69881	97.23973
Cd (Kd)	4.357508	0.975351	17.54193
Pb (Kd)	24.21178	42.50396	57.58473
As (Kd)	0.991714	1.04887	3.396427
Hg (Kd)	2.092691	1.096519	27.07087
Cu (Kd)	13.36993	26.9357	13.36993
Zn (Kd)	19.35103	32.6185	24.35103

### 3.3.3 Self-Organizing Map (SOM)

A Self-Organizing Map (SOM) methodology was employed in this study to delineate the spatial distribution of heavy metals across a dataset of 150 soil samples, each rigorously analyzed for heavy metal concentration. The SOM was iteratively trained on the data of various heavy metals, culminating in developing an intricate weight matrix for each element. The output of the SOM is illustrated in Figure 3.7 (A), where the parameters of each neuron are represented via a colour gradient, with red indicating higher concentrations and blue demonstrating lower levels. The SOM design presents a structured grid that displays the distribution of heavy metals, making it easier to understand and more precise to visualize.



Each hexagon on this grid is accurately colour-coded to represent the concentration of metals, creating an informative and detailed representation.

Significantly, the SOM results demonstrate a spatial heterogeneity in heavy metal concentrations. Heavy metals such as Chromium (Cr), Cadmium (Cd), Lead (Pb), Arsenic (As), and Nickel (Ni) predominantly manifest in the right section of the grid, indicating elevated concentrations in these areas. Conversely, Copper (Cu), Zinc (Zn), and Mercury (Hg) show higher concentrations in the upper region of the grid.

Moreover, the grid elucidates a distinct clustering pattern, mapping into three clusters (Figure 3.7 B). This pattern highlights the heterogeneity in heavy metal concentrations across the soil samples. Cluster I, characterized by the lowest heavy metal concentrations, is likely representative of the upper portion of the stream's soil samples, indicating minimal contamination. Cluster II, exhibiting the highest levels of contamination, aligns with the mid-section of the stream, a segment potentially subjected to increased environmental disturbance. Cluster III, displaying moderate contamination levels, presumably represents the downstream dispersal of heavy metals in the soil.



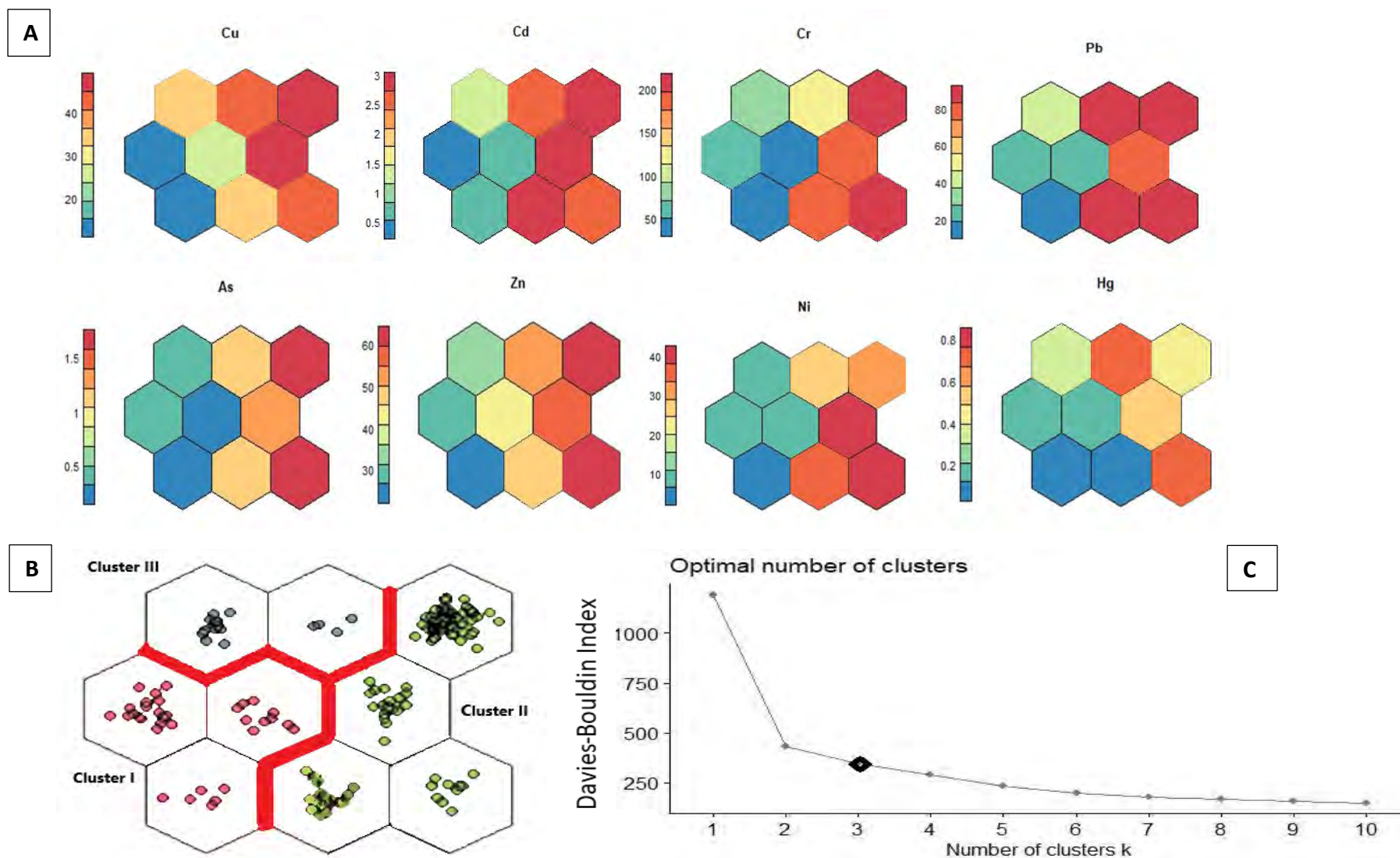


Figure 3. 7 This visualization shows the content of soil elements using a SOM (Self Organizing Map) technique. The numbers indicate the values of variables, and each neuron parameter value is color-coded. Higher concentrations of heavy metals are indicated in red, while lower concentrations are shown in blue. (B) Sample mapping and SOM clustering (C) The distinctions of the Davies-Bouldin Index with the finest sum of SOM clusters.



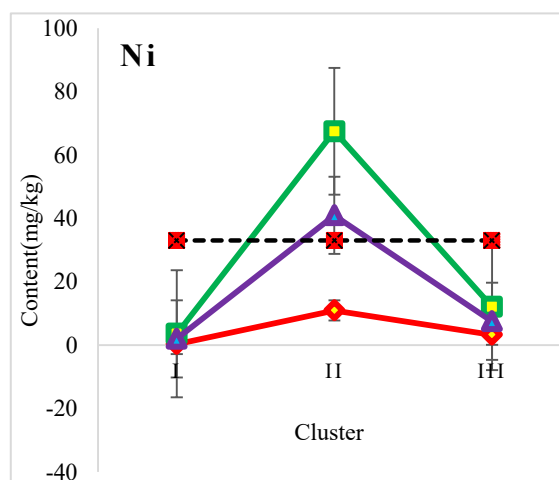
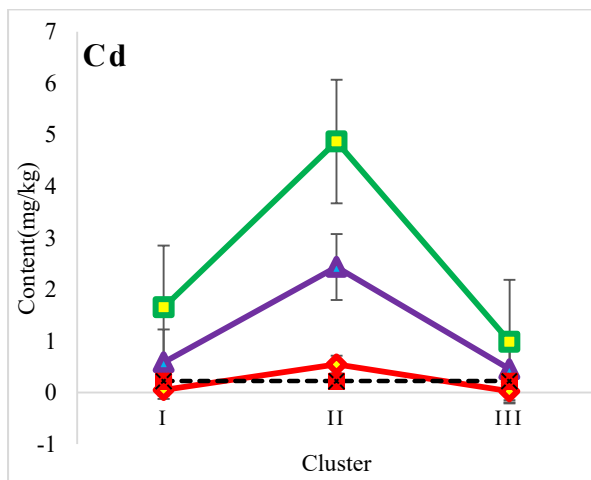
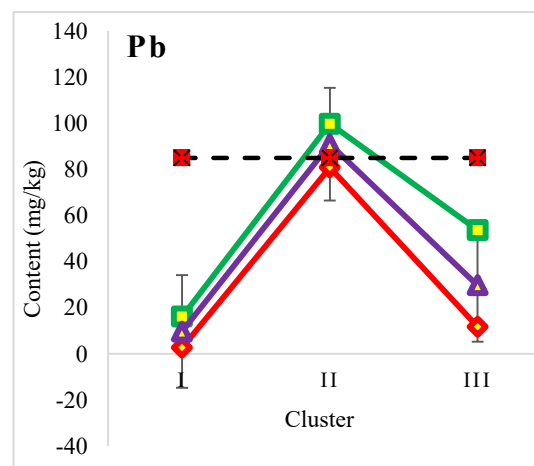
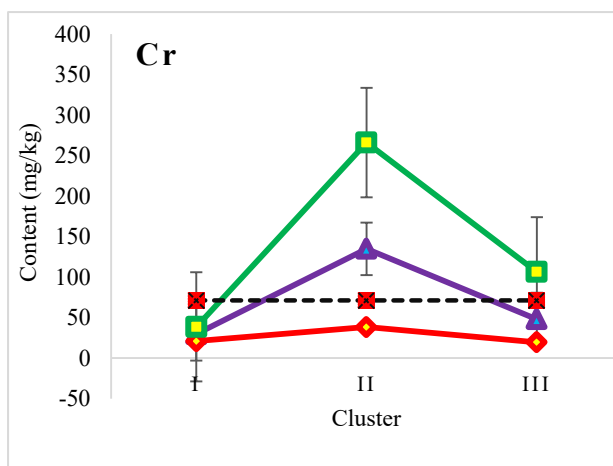
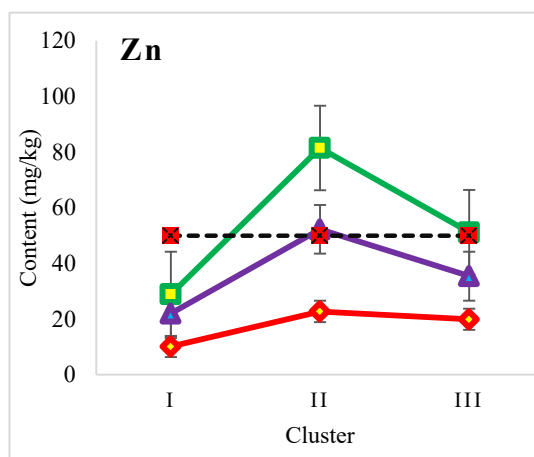
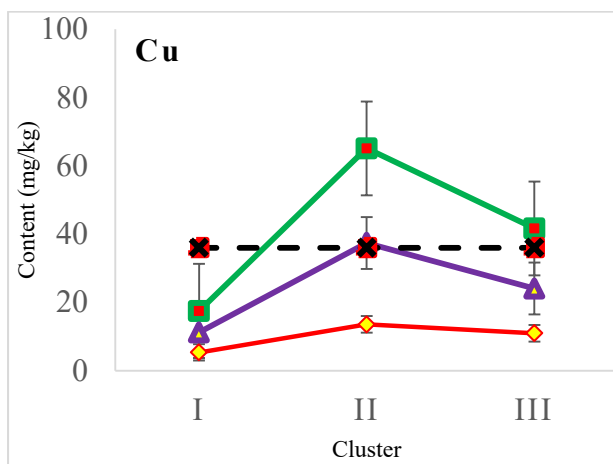
### ***3.3.4 Cluster Analysis Using Davies-Bouldin Index (DBI)***

The application of the Davies-Bouldin Index (DBI), detailed in Figure 3.7 C, allowed for an advanced clustering of the soil samples into three distinct groups based on their heavy metal concentration profiles. This clustering corroborates the distribution patterns observed in the SOM analysis and provides a nuanced classification of the soil samples, offering insights into the patterns of heavy metal accumulation in soils.

Analyzing soil samples made on the Self-Organizing Map (SOM) results, we observed sizeable variation in heavy metal contents across three distinct clusters (Figure 3.8). Cluster I, comprising 51 samples, predominantly exhibited heavy metal concentrations within established permissible limits, suggesting a relatively lower contamination profile. However, an exception was observed in the Cadmium (Cd) situation, where the maximum recorded concentration was 1.652 mg/kg, above the permissible threshold of 0.8 mg/kg. This anomaly indicates a potential for localized contamination in this cluster.

In contrast, Cluster II, consisting of 66 samples, demonstrated a markedly different contamination level. The maximum concentrations of several heavy metals, including Chromium (Cr) at 266.06 mg/kg, Nickel (Ni) at 67.44 mg/kg, Cadmium (Cd) at 4.86 mg/kg, Zinc (Zn) at 81.51 mg/kg, Copper (Cu) at 65.12 mg/kg, and Lead (Pb) at 129.75 mg/kg, were all observed to exceed the permissible limits. This data strongly suggests a significant level of contamination across multiple elements within Cluster II. Interestingly, Arsenic (As) Mercury and (Hg) did not follow this trend, with their average contents remaining within safe limits at 1.98 mg/kg and 0.92 mg/kg, respectively, indicating that contamination from these elements was not a prevalent concern in these samples. Cluster III, consisting of 33 samples, exhibited a range of concentrations for heavy metals. Notably, the maximum values for Copper (Cu) at 41.67 mg/kg, Zinc (Zn) at 51.25 mg/kg, and Chromium (Cr) at 106.41 mg/kg were found to surpass permissible limits, signaling potential contamination concerns. However, the peak amount of other heavy metals such as Lead (Pb) at 53.64 mg/kg, Nickel (Ni) at 12.11 mg/kg, Arsenic (As) at 0.563 mg/kg, and Mercury (Hg) at 0.334 mg/kg, were all recorded below their respective permissible thresholds, indicating a less pronounced contamination profile for these elements in Cluster III. These findings underscore the heterogeneous nature of heavy metal contamination across different sample clusters and associated river sections, highlighting the necessity for targeted environmental monitoring and remediation efforts tailored to specific contamination profiles.







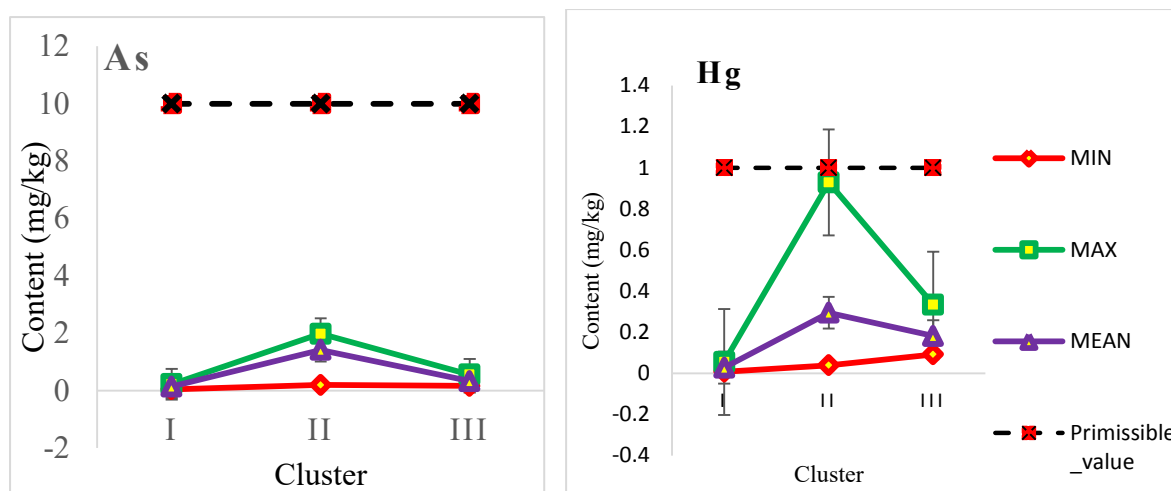


Figure 3. 8 Minimum, mean, and maximum values of soil heavy metal elements in different clusters.

### 3.3.5 Ecological risks assessment of heavy metals in soil

Ecological risks caused by heavy metals in the soil were conducted using the Potential Ecological Risk Index (PERI). The findings identified varied broadly, ranging from 11.39 to 346.01 (Table 3.5 and Appendix 4), indicating a diverse impact of heavy metals in the study area. The results indicated that most sites, about 52%, exhibited potential ecological risks, with PERI values between 150 and 300. A proportion, 8% of the total sites, demonstrated very high ecological risks, with PERI values ranging from 300 to 600. Conversely, about 40% of the sites had low ecological risks, with PERI values falling below 150.

Cadmium (Cd) was identified as the main ecological risk factor. It accounted for over 94% of the ecological risks at 69 sampling sites and contributed 64% to the total ecological risk index of the area. The ecological risks associated with Cd ranged between 3.24 and 258.5. Lead (Pb) was a significant contributor, up to 17.2% to the total PERI and risk values between 1.74 and 66.07. Chromium (Cr) also played a notable role in ecological risk, contributing 6.11% to the total PERI, with risk values ranging from 1.504 to 20.4. Other metals such as Cu, Ni, Zn, Hg and As contributed to varying extents, with Cu at 5.68%, Ni at 5.31%, Zn at 0.64%, Hg at 0.56%, and As at 0.31%. The potential ecological risks (PERI) of heavy metals across the clusters given by SOM show that Cluster II, which emerged as the highest risk area, shows the highest ecological risk of heavy metals in soil (PERI=13790.84) (Figure 3.9).

In contrast, cluster I exhibits the lowest ecological risk of heavy metals. Cluster III presents a restrained level of ecological risk. This cluster exhibits a unique set of risk factors



compared to the other two, highlighting the diverse impact of environmental and human actions on ecological risks in different areas (Figure 3.9) and (Table 3.5).

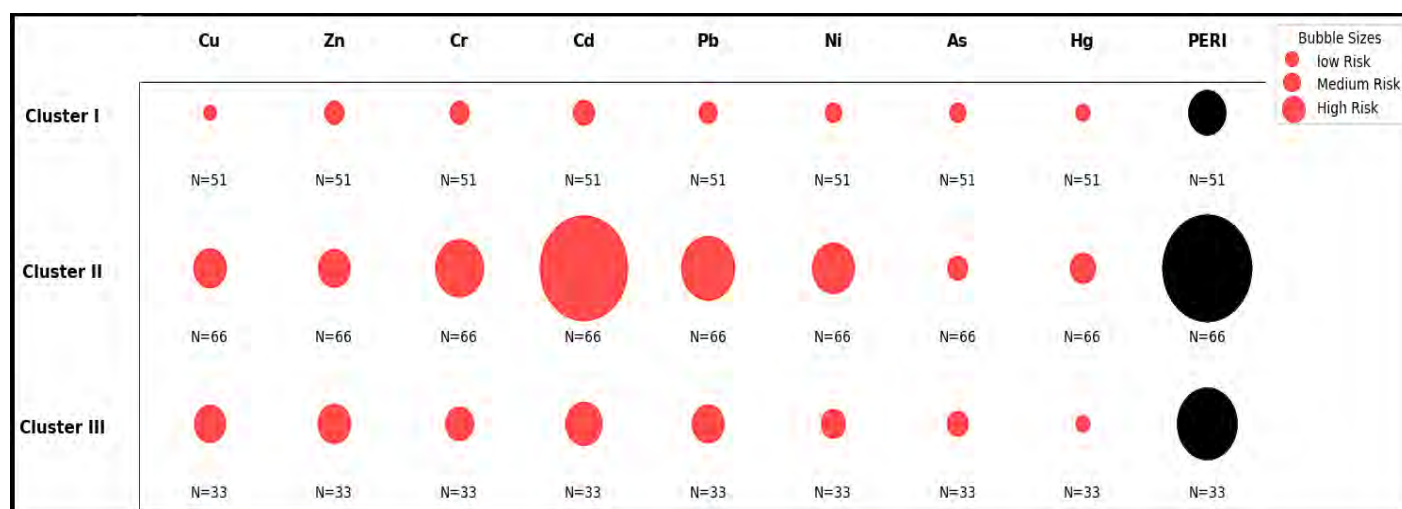
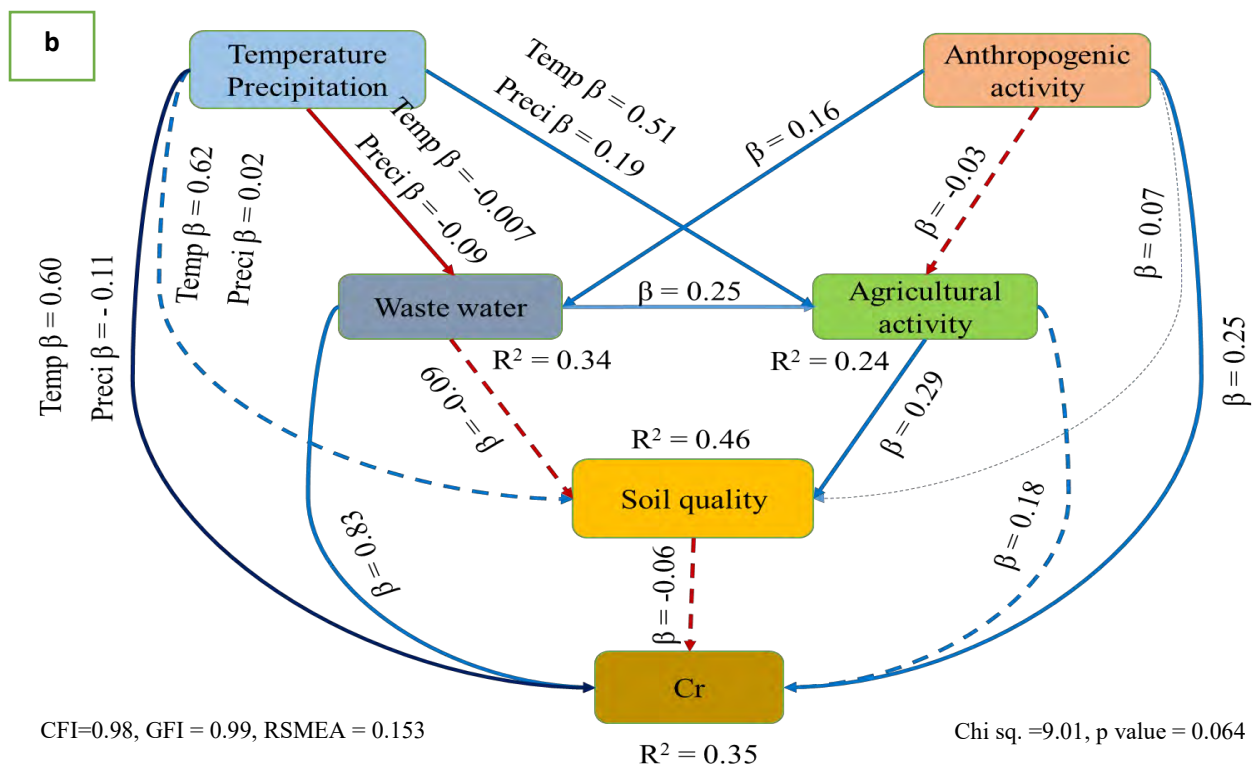
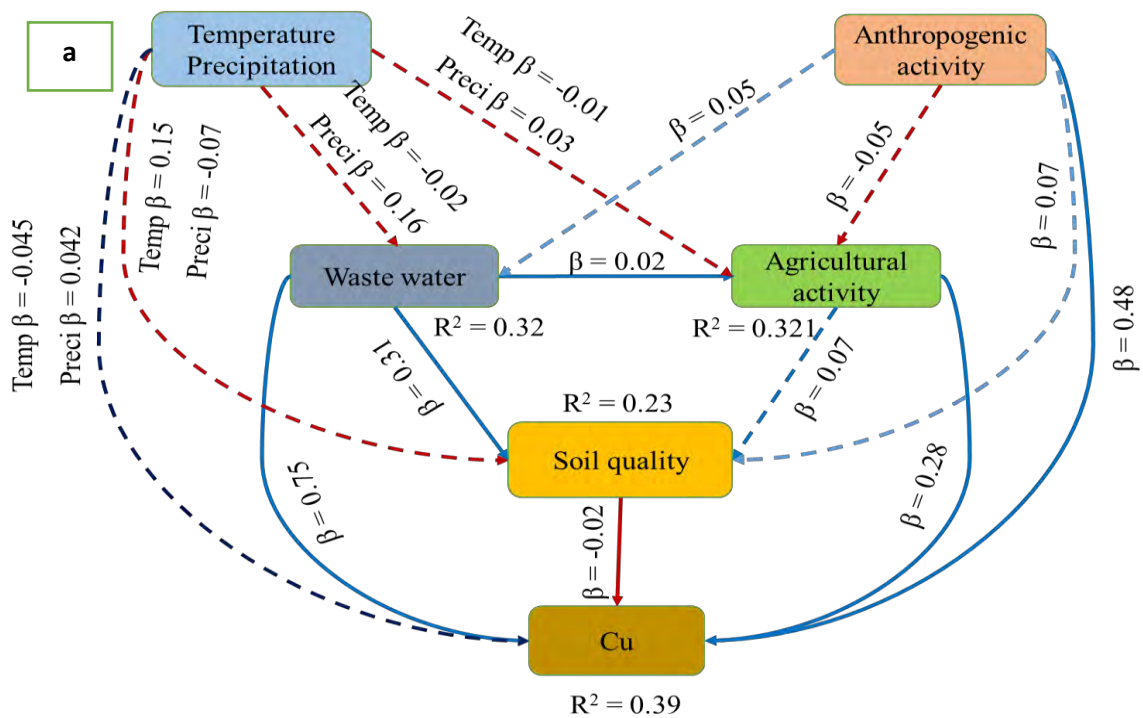


Figure 3. 9 Assessment of ecological risk from various heavy metals, both individually and within three distinct SOM clusters, in the Aik-Stream Region of Sialkot, Pakistan

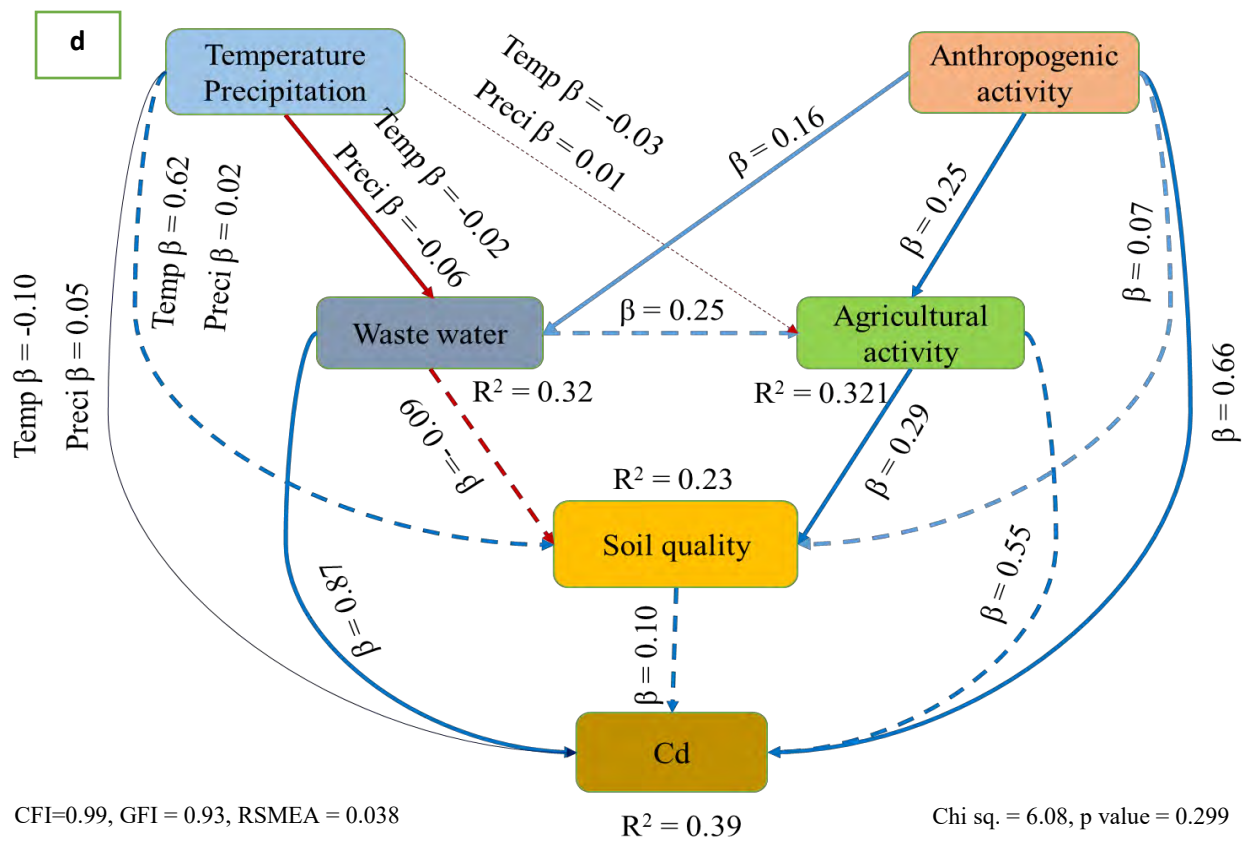
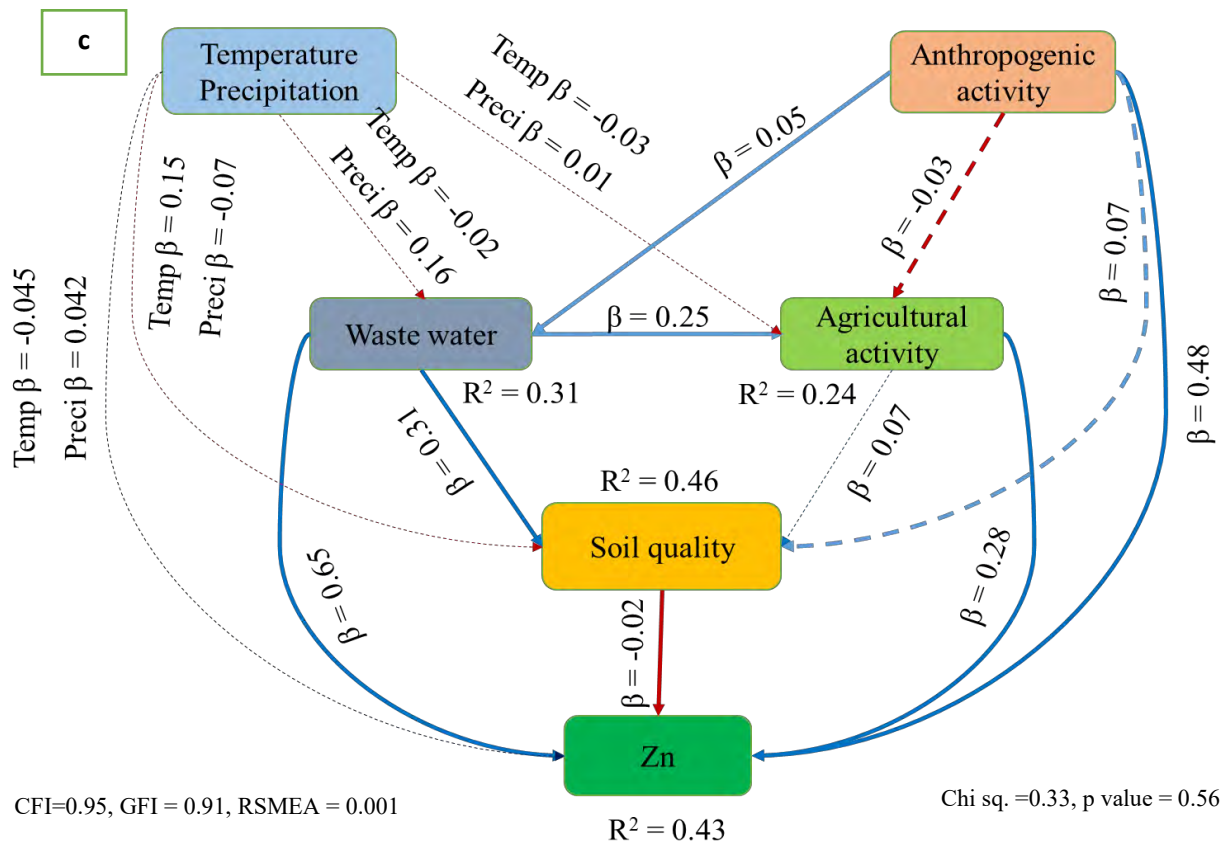
### 3.3.6 Identifying Key Factors Influencing Heavy Metals' Ecological Risk via Structural Equation Modeling

The ecological risk posed by heavy metals (Cu, Cd, Zn, Cr, Pb, Ni, As, Hg) in the study area is inclined by several factors, such as the release of industrial wastewater, soil quality, agricultural practices, human activities and weather conditions. Our objective was to comprehend how these factors interact and contribute to the risk caused by these heavy metals in soil. We used structural equation modeling (SEM) as an analysis tool to test our hypothesis. A significant positive relationship was particularly observed with increased heavy metal concentrations correlating with industrial wastewater discharge. This was evidenced by elevated levels of Cr ( $\beta$  0.83,  $p$  0.008), Pb ( $\beta$  0.89,  $p$  0.001), Cu ( $\beta$  0.75,  $p$  0.007), Cd ( $\beta$  0.87,  $p$  0.002), Ni ( $\beta$  0.66,  $p$  0.001), and Zn ( $\beta$  0.65,  $p$  0.003), as elaborated in Figure 8(a-h). Additionally, agricultural activities significantly influenced the increase in soil concentrations of Zn ( $\beta$  0.28,  $p$  0.001) and Cu ( $\beta$  0.28,  $p$  0.04), though no significant associations were noted with Cd, Cr, As, and Ni. Other variables, such as climatic conditions (temperature and precipitation), anthropogenic pressures, and diverse soil properties (pH, EC, and TDS), were also identified as having a significant direct and indirect influence on the escalation of heavy metals in soil. While not directly influencing heavy metal levels, these factors appear to exert their effects through intermediate processes. The comprehensive outcomes of the SEM analysis are detailed in Figure 3.10 (a-h) and Appendix Table 5 to 20).

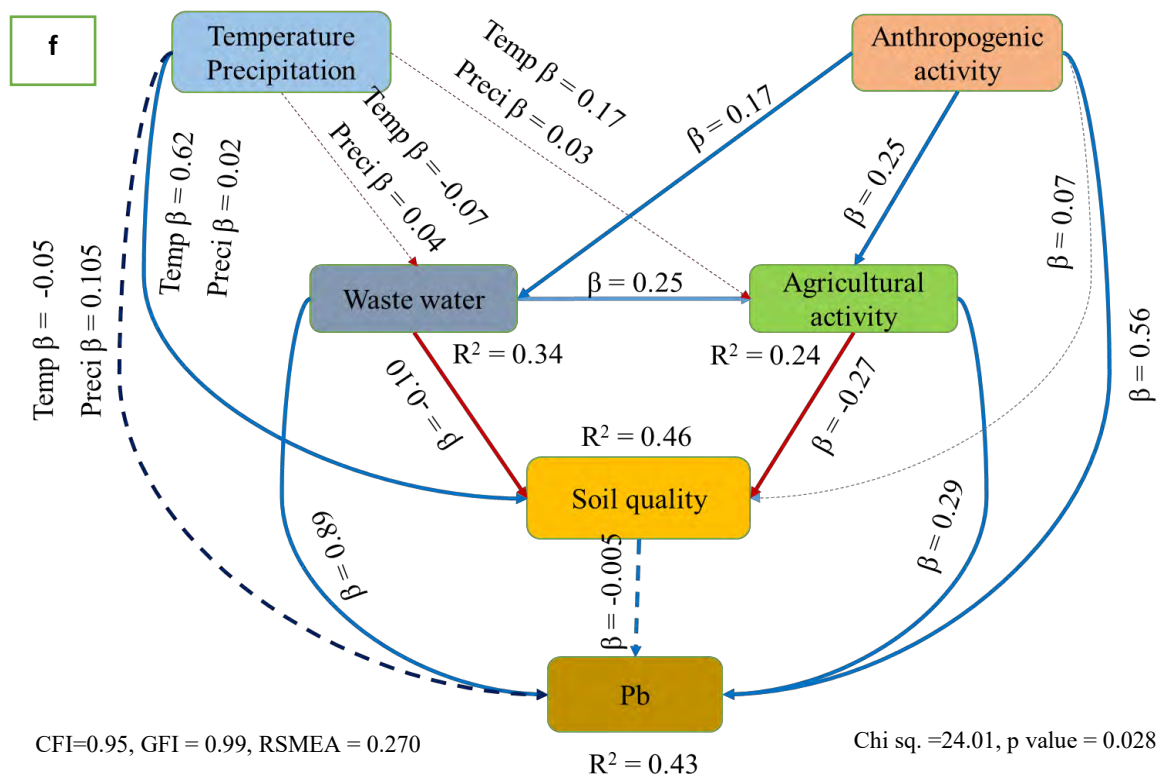
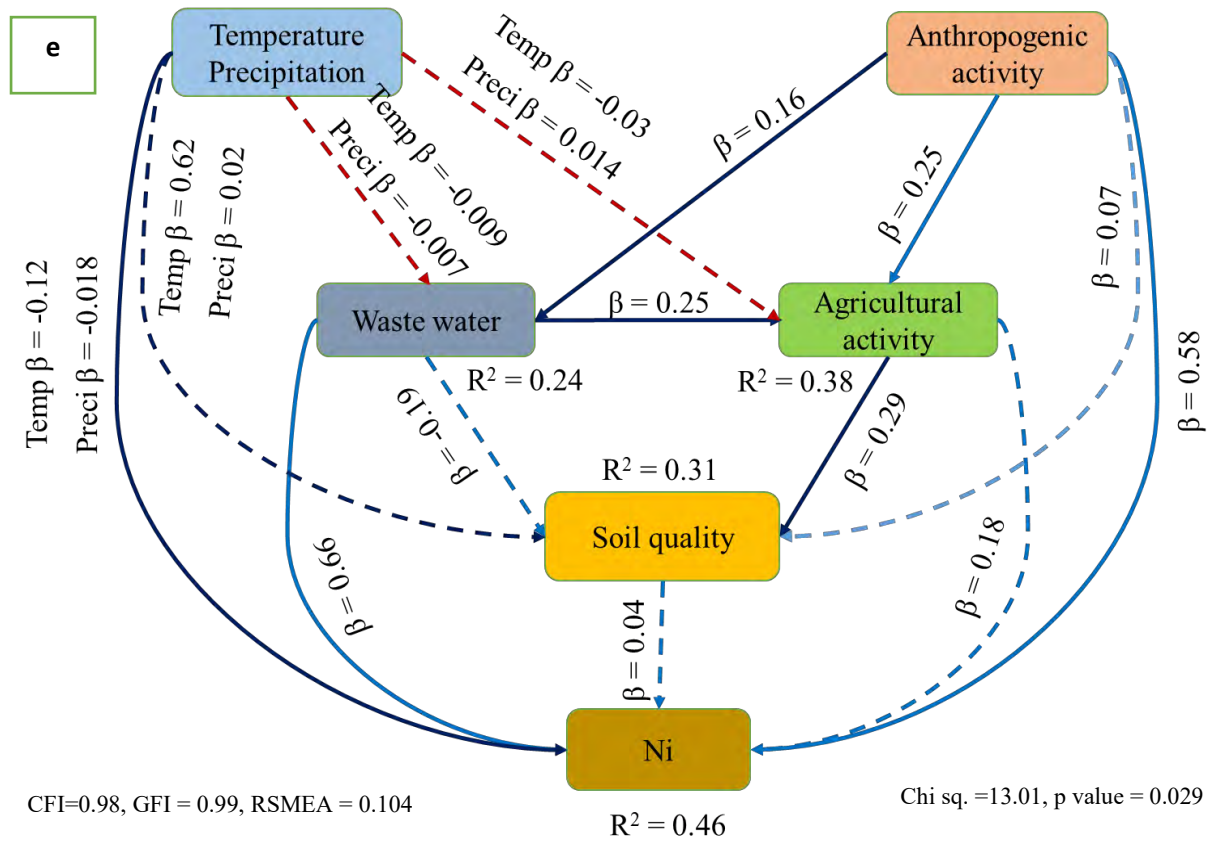














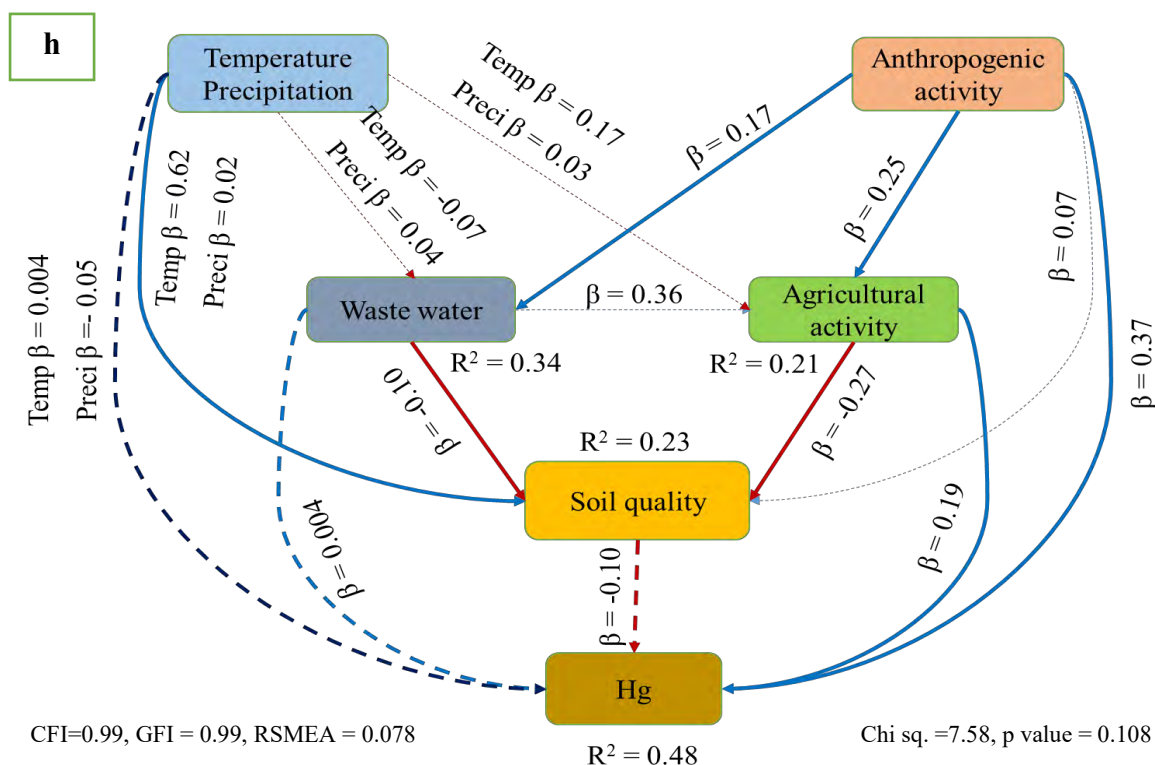
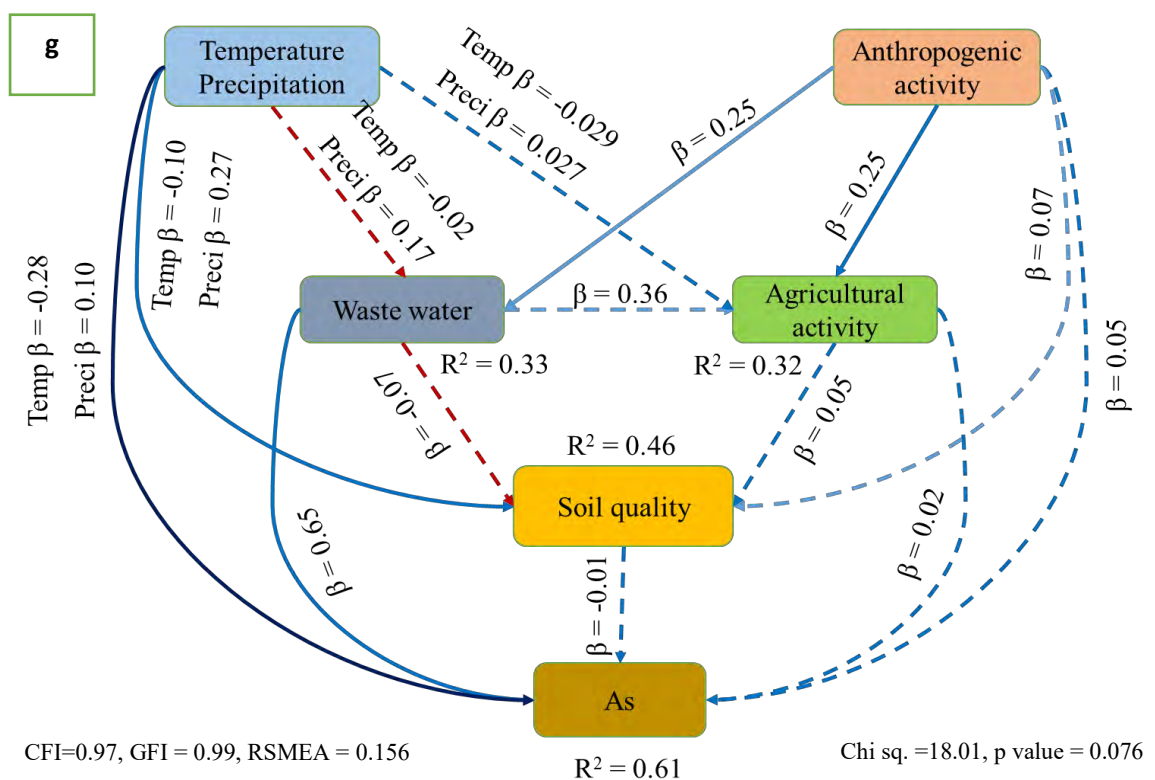


Figure 3.10 (a-h) Illustrations depict a linear structural equation model connecting various soil heavy metals, comprising Cu, Cr, Cd, Pb, Zn, Ni, As, and Hg, with factors such as soil quality, water quality, agricultural activity, and anthropogenic pressure within the Aik stream study area. Bold red and blue lines highlight significant relationships. Blue lines



represent positive and statistically significant effects, while red lines signify negative and statistically significant effects. Lines represent insignificant impacts among the variables. Red lines indicate significant relationships, while blue and dotted lines suggest statistically insignificant connections.

Table 3.4: Contamination Factor (CF) and Degree of Contamination (DC) of heavy metal in the soil around the Aik-stream.

S. S	Contamination Factor (CF)								DC
	Cu	Zn	Cr	Pb	Cd	Ni	As	Hg	
S1	0.149	0.223	0.254	0.144	0.803	0.078	0.007	0.008	1.665
S2	0.148	0.203	0.226	0.179	0.643	0.092	0.005	0.012	1.509
S3	0.287	0.443	0.385	0.120	0.545	0.085	0.009	0.017	1.892
S4	0.237	0.473	0.216	0.143	0.890	0.066	0.007	0.021	2.053
S5	0.182	0.433	0.225	0.131	0.803	0.076	0.009	0.021	1.879
S6	0.237	0.374	0.346	0.109	1.053	0.076	0.010	0.022	2.226
S7	0.149	0.520	0.330	0.099	0.068	0.060	0.011	0.012	1.248
S8	0.288	0.259	0.279	0.052	0.080	0.089	0.011	0.041	1.100
S9	0.233	0.299	0.288	0.099	0.939	0.096	0.014	0.055	2.023
S10	0.312	0.233	0.270	0.075	0.264	0.058	0.015	0.018	1.244
S11	0.284	0.214	0.244	0.156	0.680	0.047	0.016	0.034	1.675
S12	0.395	0.492	0.306	0.125	0.443	0.044	0.020	0.038	1.863
S13	0.458	0.437	0.320	0.130	0.138	0.067	0.019	0.034	1.603
S14	0.486	0.377	0.289	0.031	0.105	0.039	0.022	0.024	1.373
S15	0.429	0.377	0.336	0.077	2.065	0.018	0.023	0.045	3.371
S16	0.379	0.548	0.339	0.107	1.651	0.028	0.020	0.031	3.103
S17	0.323	0.468	0.308	0.190	0.810	0.047	0.007	0.008	2.161
S18	0.240	0.487	0.289	0.180	1.193	0.010	0.005	0.012	2.415
S19	0.405	0.559	0.317	0.103	0.771	0.015	0.009	0.017	2.196
S20	0.405	0.579	0.294	0.174	2.053	0.045	0.007	0.021	3.576
S21	0.348	0.435	0.306	0.137	1.781	0.016	0.009	0.021	3.054
S22	0.327	0.409	0.280	0.090	0.105	0.032	0.010	0.022	1.275
S23	0.326	0.350	0.246	0.102	0.776	0.039	0.011	0.012	1.860
S24	0.327	0.488	0.265	0.074	1.068	0.038	0.011	0.041	2.313
S25	0.302	0.473	0.313	0.059	0.765	0.102	0.014	0.055	2.082
S26	0.357	0.433	0.354	0.094	0.195	0.102	0.015	0.018	1.568
S27	0.329	0.489	0.323	0.080	0.926	0.046	0.016	0.034	2.244
S28	0.273	0.435	0.369	0.161	1.051	0.081	0.020	0.038	2.428
S29	0.329	0.468	0.350	0.137	0.318	0.052	0.019	0.034	1.707
S30	0.329	0.488	0.378	0.113	0.818	0.068	0.022	0.024	2.240
S31	0.304	0.513	0.326	0.095	1.234	0.010	0.023	0.045	2.550
S32	0.415	0.512	0.246	0.083	0.710	0.067	0.020	0.031	2.086
S33	0.248	0.395	0.235	0.071	1.155	0.068	0.007	0.008	2.187
S34	0.230	0.489	0.222	0.106	0.289	0.092	0.005	0.012	1.444
S35	0.312	0.569	0.231	0.070	0.193	0.076	0.009	0.017	1.477
S36	0.340	0.490	0.242	0.127	0.063	0.044	0.007	0.021	1.333
S37	0.185	0.553	0.208	0.137	1.156	0.038	0.009	0.021	2.308
S38	0.408	0.395	0.287	0.172	0.306	0.039	0.010	0.022	1.639
S39	0.351	0.512	0.258	0.110	0.193	0.056	0.011	0.012	1.503



S40	0.285	0.447	0.226	0.109	0.943	0.094	0.011	0.041	2.156
S41	0.257	0.428	0.305	0.087	0.803	0.062	0.014	0.055	2.011
S42	0.257	0.487	0.275	0.096	1.089	0.056	0.015	0.018	2.293
S43	0.308	0.412	0.357	0.049	0.780	0.079	0.016	0.034	2.036
S44	0.365	0.434	0.337	0.049	0.306	0.056	0.020	0.038	1.605
S45	0.393	0.393	0.366	0.072	0.803	0.044	0.019	0.034	2.124
S46	0.307	0.348	0.307	0.101	0.056	0.038	0.022	0.024	1.203
S47	0.335	0.488	0.375	0.149	0.068	0.084	0.023	0.045	1.566
S48	0.279	0.547	0.249	0.186	0.193	0.052	0.020	0.031	1.558
S49	0.326	0.538	0.304	0.158	0.818	0.057	0.019	0.024	2.243
S50	0.326	0.498	0.313	0.157	0.803	0.041	0.022	0.045	2.205
S51	0.346	0.500	0.317	0.158	1.431	0.047	0.020	0.039	2.858
S52	0.376	0.467	0.469	1.022	0.678	1.221	0.089	0.094	4.414
S53	0.376	0.666	0.879	1.064	1.951	1.204	0.096	0.095	6.331
S54	0.404	0.628	0.668	1.063	1.431	1.261	0.089	0.074	5.618
S55	0.583	0.475	0.718	1.064	2.938	1.234	0.115	0.928	8.054
S56	0.387	0.455	0.656	1.050	3.306	1.240	0.157	0.125	7.377
S57	0.593	0.596	0.597	1.028	3.315	1.229	0.111	0.181	7.649
S58	0.655	0.687	0.424	1.014	1.983	1.145	0.138	0.097	6.143
S59	0.722	0.627	0.384	1.013	2.820	1.173	0.198	0.356	7.294
S60	0.798	0.608	0.524	1.128	2.051	1.173	0.164	0.446	6.892
S61	0.903	0.667	0.507	1.060	2.434	1.260	0.172	0.385	7.389
S62	0.903	0.787	0.516	1.172	2.938	1.236	0.195	0.172	7.919
S63	0.655	0.707	0.488	1.174	4.314	1.520	0.164	0.542	9.563
S64	0.648	0.820	0.493	1.164	3.574	1.202	0.184	0.342	8.427
S65	0.606	0.815	0.514	1.003	2.928	1.179	0.197	0.245	7.486
S66	0.704	0.914	0.382	1.031	2.801	1.144	0.175	0.164	7.316
S67	0.657	0.835	0.430	1.052	2.960	0.972	0.183	0.195	7.284
S68	0.879	0.835	0.509	0.952	3.943	1.173	0.134	0.172	8.597
S69	0.735	0.915	0.558	1.029	2.793	1.153	0.164	0.194	7.541
S70	0.895	1.230	1.618	1.084	2.343	1.178	0.167	0.416	8.930
S71	0.618	1.284	1.596	1.083	2.808	1.299	0.089	0.368	9.145
S72	0.618	1.246	0.815	1.090	3.615	1.167	0.123	0.782	9.456
S73	1.130	1.113	1.236	1.133	4.196	1.209	0.020	0.039	10.077
S74	1.095	0.712	0.936	1.079	2.931	1.404	0.089	0.094	8.340
S75	1.187	0.950	1.144	1.079	3.101	1.373	0.096	0.095	9.026
S76	1.193	0.816	0.860	1.100	2.724	1.344	0.089	0.074	8.199
S77	1.248	0.999	1.760	1.113	1.568	1.382	0.115	0.928	9.112
S78	1.109	0.915	1.058	1.078	3.026	1.333	0.157	0.125	8.800
S79	1.041	1.102	1.252	1.090	2.300	1.230	0.111	0.181	8.307
S80	1.040	1.023	1.752	1.096	2.808	1.271	0.138	0.097	9.225
S81	1.067	1.084	1.554	1.086	1.651	1.202	0.198	0.356	8.199
S82	1.075	1.099	1.061	1.116	3.060	1.151	0.164	0.446	9.172
S83	1.048	1.149	1.561	1.102	4.453	1.219	0.172	0.385	11.087
S84	1.076	1.113	1.662	1.084	4.195	1.204	0.195	0.172	10.701
S85	0.932	1.087	1.774	1.117	5.845	0.989	0.164	0.542	12.452
S86	1.014	0.966	1.545	1.087	4.018	1.286	0.184	0.342	10.442
S87	1.294	0.943	1.473	1.099	2.818	1.144	0.197	0.245	9.213



<b>S88</b>	0.763	1.229	1.007	1.086	6.084	1.196	0.175	0.164	11.704
<b>S89</b>	0.795	1.215	1.018	1.077	3.314	1.293	0.183	0.195	9.090
<b>S90</b>	1.188	1.206	1.295	1.082	4.426	1.293	0.134	0.172	10.795
<b>S91</b>	1.228	1.189	1.216	1.090	3.176	1.144	0.164	0.194	9.402
<b>S92</b>	1.258	1.168	1.354	1.058	4.051	1.245	0.167	0.416	10.718
<b>S93</b>	1.201	1.085	1.346	1.090	4.018	1.218	0.089	0.368	10.415
<b>S94</b>	1.206	1.310	1.354	1.015	4.428	1.240	0.123	0.782	11.457
<b>S95</b>	1.207	1.290	1.449	1.004	2.818	1.246	0.020	0.039	9.073
<b>S96</b>	0.938	1.344	1.223	1.017	5.401	1.265	0.089	0.094	11.370
<b>S97</b>	1.238	1.355	1.297	1.117	5.439	1.232	0.096	0.095	11.868
<b>S98</b>	1.237	1.257	1.295	1.105	4.064	1.262	0.089	0.074	10.383
<b>S99</b>	1.315	1.272	1.194	1.090	2.806	1.241	0.115	0.928	9.961
<b>S100</b>	1.209	1.009	2.064	1.018	5.406	1.382	0.157	0.125	12.370
<b>S101</b>	1.211	1.115	2.154	1.015	2.945	1.350	0.111	0.181	10.081
<b>S102</b>	1.187	1.088	2.357	1.003	1.568	1.304	0.138	0.097	8.743
<b>S103</b>	1.174	1.548	2.661	1.016	3.103	1.156	0.198	0.356	11.211
<b>S104</b>	1.174	1.453	2.443	1.090	5.444	1.502	0.164	0.446	13.717
<b>S105</b>	1.730	1.476	2.432	1.114	4.433	1.276	0.172	0.385	13.017
<b>S106</b>	1.713	1.609	2.055	1.052	3.318	1.254	0.195	0.172	11.367
<b>S107</b>	1.435	1.630	2.116	0.964	1.695	1.276	0.164	0.542	9.823
<b>S108</b>	1.436	1.211	2.238	1.019	1.519	1.337	0.184	0.342	9.286
<b>S109</b>	1.543	1.304	2.438	1.151	1.276	1.927	0.197	0.245	10.082
<b>S110</b>	1.460	1.303	1.427	0.950	1.443	1.328	0.175	0.164	8.250
<b>S111</b>	1.212	1.508	2.415	1.090	2.065	0.378	0.183	0.195	9.047
<b>S112</b>	1.239	1.025	2.553	1.095	1.280	0.492	0.134	0.172	7.991
<b>S113</b>	1.241	0.976	2.547	1.090	1.295	0.610	0.164	0.194	8.118
<b>S114</b>	1.684	1.209	1.519	1.102	1.318	0.550	0.167	0.416	7.964
<b>S115</b>	1.809	1.113	2.448	1.039	2.140	0.371	0.089	0.368	9.377
<b>S116</b>	1.269	0.892	1.443	0.992	1.815	0.313	0.123	0.782	7.629
<b>S117</b>	0.984	1.270	1.401	1.129	2.064	0.356	0.023	0.251	7.478
<b>S118</b>	0.871	0.843	1.006	0.471	0.106	0.298	0.037	0.132	3.765
<b>S119</b>	0.872	0.849	1.064	0.490	0.068	0.263	0.026	0.216	3.849
<b>S120</b>	0.926	0.507	0.885	0.549	0.081	0.239	0.042	0.147	3.376
<b>S121</b>	1.073	1.021	0.827	0.604	1.215	0.336	0.016	0.334	5.428
<b>S122</b>	1.158	0.571	0.737	0.596	0.053	0.267	0.024	0.146	3.551
<b>S123</b>	1.156	0.593	0.419	0.631	0.565	0.212	0.031	0.092	3.700
<b>S124</b>	0.796	0.413	0.729	0.480	0.068	0.116	0.023	0.125	2.750
<b>S125</b>	0.824	0.553	0.338	0.556	0.046	0.150	0.028	0.229	2.723
<b>S126</b>	0.767	0.655	0.726	0.502	0.349	0.148	0.056	0.122	3.325
<b>S127</b>	0.882	0.453	0.559	0.437	0.098	0.235	0.052	0.224	2.940
<b>S128</b>	0.882	0.935	0.643	0.396	0.855	0.229	0.023	0.251	4.214
<b>S129</b>	0.939	0.957	0.229	0.397	1.205	0.218	0.037	0.132	4.116
<b>S130</b>	0.610	0.813	0.346	0.246	0.056	0.114	0.026	0.216	2.427
<b>S131</b>	0.609	1.025	0.427	0.243	0.059	0.124	0.042	0.147	2.677
<b>S132</b>	0.554	1.012	0.332	0.257	1.178	0.133	0.016	0.334	3.816
<b>S133</b>	0.512	0.893	0.230	0.169	1.089	0.165	0.024	0.146	3.228
<b>S134</b>	0.485	0.891	0.520	0.331	0.955	0.301	0.031	0.092	3.606
<b>S135</b>	0.629	0.740	0.406	0.364	0.268	0.296	0.023	0.125	2.852



<b>S136</b>	0.709	0.653	0.440	0.392	0.928	0.346	0.028	0.229	3.725
<b>S137</b>	0.346	0.657	0.420	0.321	0.405	0.343	0.056	0.122	2.671
<b>S138</b>	0.567	0.782	0.319	0.219	1.231	0.284	0.052	0.224	3.680
<b>S139</b>	0.592	0.681	0.221	0.174	0.026	0.225	0.023	0.251	2.195
<b>S140</b>	0.537	0.621	0.710	0.162	0.943	0.204	0.037	0.132	3.346
<b>S141</b>	0.435	0.813	0.510	0.267	0.080	0.161	0.026	0.216	2.509
<b>S142</b>	0.574	0.830	0.326	0.225	1.105	0.121	0.042	0.147	3.370
<b>S143</b>	0.657	0.887	0.313	0.377	0.431	0.118	0.016	0.334	3.135
<b>S144</b>	0.574	0.679	0.406	0.149	1.193	0.175	0.024	0.146	3.346
<b>S145</b>	0.421	0.511	0.195	0.205	0.146	0.094	0.031	0.092	1.696
<b>S146</b>	0.616	0.533	0.372	0.375	0.818	0.235	0.023	0.125	3.096
<b>S147</b>	0.426	0.568	0.381	0.302	1.228	0.322	0.028	0.229	3.483
<b>S148</b>	0.416	0.447	0.226	0.291	0.089	0.283	0.056	0.122	1.930
<b>S149</b>	0.304	0.399	0.316	0.209	0.981	0.158	0.052	0.224	2.644
<b>S150</b>	0.326	0.613	0.251	0.138	0.564	0.161	0.056	0.122	2.231

Table 3.5 Potential Ecological Risk (PER) and Potential Ecological Risk Index (PERI) calculation of the soil around the Aik Stream.

	<b>Potential Ecological Risk (PER)</b>								
<b>S. S</b>	<b>Cu</b>	<b>Zn</b>	<b>Cr</b>	<b>Cd</b>	<b>Pb</b>	<b>Ni</b>	<b>As</b>	<b>Hg</b>	<b>PERI</b>
<b>S1</b>	1.675	0.223	1.956	32.4648	3.4298	0.85	0.04	0.0304	40.669
<b>S2</b>	1.67	0.203	1.742	26.2572	4.3572	1.005	0.03	0.0456	35.31
<b>S3</b>	3.225	0.443	2.964	22.5876	3.3376	0.935	0.04	0.0648	33.597
<b>S4</b>	2.67	0.473	1.66	36.57	3.995	0.725	0.04	0.08	46.213
<b>S5</b>	2.045	0.433	1.732	33.06	3.745	0.83	0.05	0.08	41.975
<b>S6</b>	2.67	0.374	2.662	43.098	3.323	0.825	0.05	0.084	53.086
<b>S7</b>	1.68	0.52	2.538	3.2472	2.6622	0.65	0.05	0.0456	11.393
<b>S8</b>	3.24	0.259	2.148	5.082	2.987	0.97	0.06	0.156	14.902
<b>S9</b>	2.615	0.299	2.218	40.0752	4.6302	1.045	0.07	0.2096	51.162
<b>S10</b>	3.51	0.233	2.074	11.3856	2.4106	0.64	0.08	0.0688	20.402
<b>S11</b>	3.195	0.214	1.88	28.7652	4.8652	0.515	0.08	0.1296	39.644
<b>S12</b>	4.445	0.492	2.354	19.4376	4.3976	0.48	0.1	0.1448	31.851
<b>S13</b>	5.155	0.437	2.46	7.0452	4.3252	0.73	0.1	0.1296	20.382
<b>S14</b>	5.465	0.377	2.22	5.2944	1.7494	0.425	0.11	0.0912	15.732
<b>S15</b>	4.83	0.377	2.586	84.6444	3.6894	0.2	0.11	0.1712	96.608
<b>S16</b>	4.265	0.548	2.606	67.4808	3.6958	0.305	0.1	0.1184	79.119
<b>S17</b>	3.64	0.468	2.37	32.7648	4.3948	0.515	0.04	0.0304	44.223
<b>S18</b>	2.705	0.487	2.22	48.2472	4.3622	0.11	0.03	0.0456	58.207
<b>S19</b>	4.56	0.559	2.436	31.6176	2.9626	0.165	0.04	0.0648	42.405
<b>S20</b>	4.55	0.579	2.258	83.07	4.65	0.49	0.04	0.08	95.717
<b>S21</b>	3.915	0.435	2.358	72.21	3.87	0.175	0.05	0.08	83.093
<b>S22</b>	3.675	0.409	2.156	5.208	2.923	0.35	0.05	0.084	14.855
<b>S23</b>	3.665	0.35	1.89	31.5972	2.7072	0.425	0.05	0.0456	40.73
<b>S24</b>	3.675	0.488	2.038	44.562	3.452	0.42	0.06	0.156	54.851
<b>S25</b>	3.39	0.473	2.408	33.1152	3.7602	1.115	0.07	0.2096	44.541
<b>S26</b>	4.015	0.433	2.726	8.6256	2.8156	1.115	0.08	0.0688	19.879
<b>S27</b>	3.7	0.489	2.488	38.6052	3.2602	0.505	0.08	0.1296	49.257



S28	3.07	0.435	2.836	43.7976	5.1526	0.89	0.1	0.1448	56.426
S29	3.7	0.468	2.69	14.2452	4.4652	0.57	0.1	0.1296	26.368
S30	3.7	0.488	2.912	33.7944	3.4844	0.74	0.11	0.0912	45.32
S31	3.425	0.513	2.512	51.4044	4.0744	0.105	0.11	0.1712	62.315
S32	4.665	0.512	1.894	29.8308	3.1908	0.735	0.1	0.1184	41.046
S33	2.795	0.395	1.81	46.5648	1.8848	0.74	0.04	0.0304	54.26
S34	2.585	0.489	1.706	12.0972	2.7922	1.005	0.03	0.0456	20.75
S35	3.515	0.569	1.78	8.4876	2.2626	0.83	0.04	0.0648	17.549
S36	3.825	0.49	1.866	3.45	3.65	0.475	0.04	0.08	13.876
S37	2.08	0.553	1.604	47.22	3.875	0.415	0.05	0.08	55.877
S38	4.585	0.395	2.212	13.248	4.663	0.42	0.05	0.084	25.657
S39	3.95	0.512	1.988	8.2572	2.8872	0.61	0.05	0.0456	18.3
S40	3.205	0.447	1.742	39.582	4.197	1.03	0.06	0.156	50.419
S41	2.895	0.428	2.348	34.6152	4.3552	0.68	0.07	0.2096	45.601
S42	2.895	0.487	2.116	44.3856	2.8606	0.61	0.08	0.0688	53.503
S43	3.47	0.412	2.744	32.7552	2.5852	0.87	0.08	0.1296	43.046
S44	4.11	0.434	2.594	13.9776	2.7676	0.61	0.1	0.1448	24.738
S45	4.42	0.393	2.82	33.6552	3.0852	0.48	0.1	0.1296	45.083
S46	3.45	0.348	2.36	3.3444	3.2344	0.415	0.11	0.0912	13.353
S47	3.765	0.488	2.886	4.7544	5.2194	0.92	0.11	0.1712	18.314
S48	3.14	0.547	1.914	9.1308	5.3858	0.57	0.1	0.1184	20.906
S49	3.67	0.538	2.338	33.7944	4.4444	0.62	0.1	0.0912	45.596
S50	3.675	0.498	2.412	34.1544	5.3944	0.445	0.11	0.1712	46.86
S51	3.89	0.5	2.44	59.0256	5.1406	0.515	0.1	0.1488	71.76
S52	4.23	0.467	3.604	31.3908	26.0108	13.355	0.44	0.3584	79.856
S53	4.23	0.666	6.758	82.3992	26.9442	13.17	0.48	0.3616	135.009
S54	4.54	0.628	5.142	60.6192	25.9592	13.79	0.45	0.2816	111.41
S55	6.56	0.475	5.52	159.9324	65.0374	13.495	0.57	3.5352	255.125
S56	4.36	0.455	5.048	137.952	28.022	13.565	0.78	0.476	190.658
S57	6.67	0.596	4.594	140.8752	30.1102	13.445	0.55	0.6896	197.53
S58	7.37	0.687	3.266	83.7252	25.9802	12.525	0.69	0.3696	134.613
S59	8.12	0.627	2.956	129.072	37.807	12.835	0.99	1.356	193.763
S60	8.98	0.608	4.03	102.4404	44.3504	12.83	0.82	1.6992	175.758
S61	10.165	0.667	3.898	114.9468	40.1318	13.785	0.86	1.4664	185.92
S62	10.165	0.787	3.972	125.3724	32.7774	13.515	0.98	0.6552	188.224
S63	7.365	0.707	3.75	197.3376	49.7176	16.63	0.82	2.0648	278.392
S64	7.295	0.82	3.796	158.5884	40.3734	13.15	0.92	1.3032	226.246
S65	6.815	0.815	3.954	128.2932	32.5132	12.89	0.99	0.9336	187.204
S66	7.92	0.914	2.942	119.5476	29.4076	12.51	0.88	0.6248	174.746
S67	7.39	0.835	3.306	127.3284	31.2834	10.63	0.92	0.7432	182.436
S68	9.885	0.835	3.918	165.5724	28.0974	12.83	0.67	0.6552	222.463
S69	8.265	0.915	4.296	120.5604	30.7354	12.615	0.82	0.7392	178.946
S70	10.07	1.23	12.444	112.7076	42.0576	12.88	0.83	1.5848	193.804
S71	6.96	1.284	12.278	129.1092	39.8242	14.205	0.45	1.4016	205.512
S72	6.95	1.246	6.272	180.3504	58.9104	12.765	0.61	2.9792	270.083
S73	12.71	1.113	9.504	169.6356	25.8706	13.23	0.1	0.1488	232.312
S74	12.315	0.712	7.204	121.5408	27.2258	15.36	0.44	0.3584	185.156
S75	13.36	0.95	8.802	128.3892	27.2692	15.015	0.48	0.3616	194.627



S76	13.425	0.816	6.612	112.3392	26.7492	14.695	0.45	0.2816	175.368
S77	14.045	0.999	13.536	105.1224	66.0774	15.11	0.57	3.5352	218.995
S78	12.475	0.915	8.136	126.762	28.622	14.575	0.78	0.476	192.741
S79	11.705	1.102	9.632	100.2852	31.4402	13.46	0.55	0.6896	168.864
S80	11.7	1.023	13.478	116.7252	27.7252	13.9	0.69	0.3696	185.611
S81	12.005	1.084	11.952	82.332	39.357	13.14	0.99	1.356	162.216
S82	12.1	1.099	8.162	142.7904	44.1104	12.585	0.82	1.6992	223.366
S83	11.785	1.149	12.01	195.7068	41.0068	13.33	0.86	1.4664	277.314
S84	12.105	1.113	12.788	175.6524	30.8874	13.17	0.98	0.6552	247.351
S85	10.48	1.087	13.65	258.5676	48.5226	10.825	0.82	2.0648	346.017
S86	11.41	0.966	11.884	176.3484	38.7434	14.065	0.92	1.3032	255.64
S87	14.555	0.943	11.332	123.9132	34.5682	12.515	0.99	0.9336	199.75
S88	8.58	1.229	7.742	250.8576	30.5826	13.075	0.88	0.6248	313.571
S89	8.94	1.215	7.83	141.4584	31.8034	14.145	0.92	0.7432	207.055
S90	13.365	1.206	9.96	184.9224	30.8524	14.14	0.67	0.6552	255.771
S91	13.815	1.189	9.358	135.9204	32.0354	12.515	0.82	0.7392	206.392
S92	14.15	1.168	10.418	181.0776	41.5076	13.62	0.83	1.5848	264.356
S93	13.51	1.085	10.356	177.5292	39.9742	13.325	0.45	1.4016	257.631
S94	13.565	1.31	10.418	212.8404	57.3104	13.565	0.61	2.9792	312.598
S95	13.585	1.29	11.144	114.4956	23.1206	13.635	0.1	0.1488	177.519
S96	10.545	1.344	9.408	220.3608	25.9058	13.835	0.44	0.3584	282.197
S97	13.925	1.355	9.974	221.8992	28.0792	13.475	0.48	0.3616	289.549
S98	13.915	1.257	9.962	165.9192	26.8642	13.805	0.45	0.2816	232.454
S99	14.8	1.272	9.188	154.6824	65.5774	13.57	0.57	3.5352	263.195
S100	13.605	1.009	15.878	221.952	27.342	15.11	0.78	0.476	296.152
S101	13.62	1.115	16.57	126.0852	29.8402	14.765	0.55	0.6896	203.235
S102	13.355	1.088	18.134	67.1352	25.7552	14.265	0.69	0.3696	140.792
S103	13.205	1.548	20.466	140.382	37.862	12.64	0.99	1.356	228.449
S104	13.21	1.453	18.794	238.1304	43.5504	16.435	0.82	1.6992	334.092
S105	19.46	1.476	18.706	194.8968	41.2568	13.955	0.86	1.4664	292.077
S106	19.265	1.609	15.804	140.5524	30.2274	13.715	0.98	0.6552	222.808
S107	16.15	1.63	16.278	92.5776	45.2676	13.95	0.82	2.0648	188.738
S108	16.155	1.211	17.218	76.3884	37.3034	14.62	0.92	1.3032	165.119
S109	17.365	1.304	18.756	62.2632	35.6632	21.075	0.99	0.9336	158.35
S110	16.43	1.303	10.978	65.1876	27.6976	14.52	0.88	0.6248	137.621
S111	13.635	1.508	18.578	91.5084	32.0784	4.14	0.92	0.7432	163.111
S112	13.945	1.025	19.642	59.0724	31.1324	5.38	0.67	0.6552	131.522
S113	13.965	0.976	19.594	60.6804	32.0304	6.675	0.82	0.7392	135.48
S114	18.945	1.209	11.682	71.7276	42.4276	6.015	0.83	1.5848	154.421
S115	20.35	1.113	18.83	102.4092	38.9042	4.06	0.45	1.4016	187.518
S116	14.27	0.892	11.096	108.3504	56.8404	3.425	0.61	2.9792	198.463
S117	11.07	1.27	10.78	94.032	35.457	3.895	0.12	0.956	157.58
S118	9.795	0.843	7.738	10.2984	16.0534	3.265	0.19	0.5032	48.686
S119	9.81	0.849	8.186	12.5784	20.2984	2.88	0.13	0.8232	55.555
S120	10.42	0.507	6.81	9.96	18.38	2.61	0.21	0.56	49.457
S121	12.075	1.021	6.364	63.864	28.104	3.68	0.08	1.272	116.46
S122	13.025	0.571	5.672	8.772	19.332	2.92	0.12	0.556	50.968
S123	13.01	0.593	3.226	26.7948	17.6148	2.32	0.16	0.3504	64.069



<b>S124</b>	8.955	0.413	5.612	8.412	15.902	1.265	0.12	0.476	41.155
<b>S125</b>	9.265	0.553	2.596	12.324	22.279	1.64	0.14	0.872	49.669
<b>S126</b>	8.63	0.655	5.588	19.5276	16.2426	1.62	0.28	0.4648	53.008
<b>S127</b>	9.92	0.453	4.304	14.1432	19.5332	2.565	0.26	0.8536	52.032
<b>S128</b>	9.925	0.935	4.942	45.672	19.887	2.51	0.12	0.956	84.947
<b>S129</b>	10.57	0.957	1.764	54.2484	14.4784	2.39	0.19	0.5032	85.101
<b>S130</b>	6.86	0.813	2.664	12.1284	15.1034	1.245	0.13	0.8232	39.767
<b>S131</b>	6.855	1.025	3.284	9.06	11.88	1.36	0.21	0.56	34.234
<b>S132</b>	6.23	1.012	2.55	62.364	20.734	1.455	0.08	1.272	95.697
<b>S133</b>	5.76	0.893	1.768	50.232	10.272	1.805	0.12	0.556	71.406
<b>S134</b>	5.46	0.891	3.996	42.3948	11.2398	3.295	0.16	0.3504	67.787
<b>S135</b>	7.08	0.74	3.126	16.422	13.457	3.24	0.12	0.476	44.661
<b>S136</b>	7.98	0.653	3.386	47.574	18.799	3.785	0.14	0.872	83.189
<b>S137</b>	3.895	0.657	3.234	21.7776	12.3876	3.755	0.28	0.4648	46.451
<b>S138</b>	6.38	0.782	2.458	59.5032	14.8982	3.11	0.26	0.8536	88.245
<b>S139</b>	6.665	0.681	1.7	12.522	15.177	2.465	0.12	0.956	40.286
<b>S140</b>	6.045	0.621	5.464	43.7484	9.4884	2.225	0.19	0.5032	68.285
<b>S141</b>	4.89	0.813	3.92	13.0884	15.5634	1.765	0.13	0.8232	40.993
<b>S142</b>	6.455	0.83	2.51	50.91	11.49	1.325	0.21	0.56	74.29
<b>S143</b>	7.39	0.887	2.412	32.514	23.284	1.295	0.08	1.272	69.134
<b>S144</b>	6.46	0.679	3.12	54.372	9.837	1.915	0.12	0.556	77.059
<b>S145</b>	4.735	0.511	1.504	10.0548	8.5648	1.025	0.16	0.3504	26.905
<b>S146</b>	6.925	0.533	2.858	38.412	13.687	2.565	0.12	0.476	65.576
<b>S147</b>	4.795	0.568	2.928	59.574	16.879	3.515	0.14	0.872	89.271
<b>S148</b>	4.685	0.447	1.742	9.1176	11.7526	3.1	0.28	0.4648	31.589
<b>S149</b>	3.415	0.399	2.434	49.4832	14.6782	1.73	0.26	0.8536	73.253
<b>S150</b>	3.67	0.613	1.928	29.232	9.607	1.765	0.12	0.556	47.491



### 3.4 Discussion

The principal focus of this study was the comprehensive assessment of soil pollution in the area surrounding the Aik-stream, which has been subject to contamination from industrial, agricultural, and municipal wastewater discharges (Chapter 1). This pollution has led to the accumulation of heavy metals in the soil, a matter of considerable concern due to their persistence and toxicity.

To evaluate this heavy metal contamination, soil samples from the vicinity of the Aik Stream were analyzed, focusing on two pivotal indices, i.e., the Contamination Factor (CF) and the Degree of Contamination (DC). These indices, widely recognized in environmental research, provide a quantifiable measure of heavy metal contamination intensity (Bello, Zakari, Ibeanu, & Muhammad, 2015; El-Amier et al., 2017; Ravankhah et al., 2015; Said, Salman, & Elnazer, 2019) underscored the effectiveness of these indices, both individually and in combination, for evaluating soil contamination. The application of these indices to our soil samples revealed that Nickel (Ni), Cadmium (Cd), Lead (Pb), and Chromium (Cr) are the predominant contributors to soil pollution in the area. Meanwhile, Zinc (Zn) and Copper (Cu) were found to moderately pollute the soil, with Arsenic (As) and Mercury (Hg) contributing minimally. Notably, the maximum contamination degree near the Aik-stream was determined to be 809.69, a value that markedly surpasses those reported in similar studies, such as the one conducted by (Malik, Jadoon, et al., 2010) in different soil regions of the Sialkot Industrial zone. Through applying these indices, we gained substantial insights into the patterns and levels of heavy metal contamination, especially in areas with concentrated heavy metal deposits, notably in the mid-stream region of Aik-stream. This region, subjected to considerable quantities of wastewater and urban waste, demonstrates the significant impact of industrial and urban processes on soil quality. The heightened concentrations of heavy metals observed in this midstream location indicate a significant influence from industrial and human activities, emphasizing the urgent need for specialized environmental management and remediation strategies in these regions.

Additionally, this study classified and analyzed soil heavy metal pollution in the Aik-Stream area using a Self Organizing Map (SOM). This method allowed for a more nuanced knowledge of the contamination patterns and levels, providing a sophisticated method for identifying and classifying the various types and degrees of pollution in the soil. This study extends the foundational approaches established by (Vesanto & Alhoniemi, 2000), particularly in environmental data visualization and complex data analysis. According to this



study, the Aik-stream area was effectively separated into three distinct zones through Self-Organizing Maps (SOM), each with its unique pollution profile and management requirements. This approach aligns with the work of (Nguyen et al., 2021), who investigated using advanced SOM techniques for handling complex environmental datasets. Their advancements in processing intricate environmental data mirror the challenges we faced in analyzing soil heavy metal pollution data, underscoring the versatility and effectiveness of SOM in environmental data analysis. The study by (Q. Wang, Jiang, Gao, Zhang, & Chang, 2022) on SOM applications in urban environmental management provides relevant insights into practical implementations in urban settings. Their findings, which closely align with our work in urban industrial areas, demonstrate the effectiveness of SOM in urban environmental planning and management strategies. Additionally, (Corona, Mulas, Baratti, & Romagnoli, 2010) and (Carlei & Nuccio, 2014) on using SOM to identify pollution sources in industrial areas complement our research. Their work underscores SOM's precision in pinpointing pollution sources, a crucial aspect of our goal to classify distinct pollution zones. Their findings and methodologies have contributed significantly to enhancing our ability to accurately identify and categorize the different sources and types of heavy metal contamination within the Aik-stream area, thereby aiding in the development of targeted management and remediation strategies. In our research, the utilization of Self-Organizing Maps (SOM) has been pivotal in dividing the study area into three distinct zones, each characterized by their own unique features and specific management requirements

In our study, the designation of Cluster I as a Low-Risk Area emphasizes the need for vigilant environmental monitoring. This finding, resonating with the recommendations of (Hull et al., 2016), challenges the conventional focus on high-risk areas alone and highlights the importance of maintaining environmental quality through consistent monitoring and potential remediation in areas perceived as less contaminated. In contrast, Cluster II, identified as a High-Risk Urban Industrial Area, brings to light the complex relationship between industrial activities and environmental health. Our recommendations for this cluster, aligning with the insights of (Grainger-Brown, Malekpour, Raven, & Taylor, 2022) and (De Jong, Joss, Schraven, Zhan, & Weijnen, 2015), promise a shift towards sustainable industrial practices and cleaner production methods, especially crucial in urban settings. Meanwhile, the classification of Cluster III as a Medium-Risk Early Warning Area underscores the effectiveness of anticipatory strategies in environmental management. Aligning with (Hobbs & Cramer, 2008) and (Alkhafaji & Nelson, 2013), this finding advocates for early



interventions and preventive measures, highlighting the importance of proactive approaches in managing ecological risks in dynamic environments. Using Self-Organizing Maps (SOM) in our study has helped categorize these distinct zones and provided more profound insights into the varying degrees of pollution, underscoring the need for tailored environmental management strategies specific to each zone.

Additionally, our study evaluates the potential ecological risks posed by heavy metals in soil using the Potential Ecological Risk Index, a methodology pioneered by (Hakanson, 1980a). The investigation into the ecological risks of heavy metals in soil exposed significant variations in risk indexes, with particular emphasis on cadmium (Cd), lead (Pb), and chromium (Cr) as critical contributors to these risks. These results align with the concerns highlighted in prior research regarding the extensive impact of these metals on environmental health and safety (Balali-Mood, Naseri, Tahergorabi, Khazdair, & Sadeghi, 2021; Z. Rahman & Singh, 2019). This underscores the critical need for ongoing research and active management strategies to address these environmental challenges.

The outcomes of our research identifying cadmium (Cd) as the primary ecological risk factor are consistent with a range of previous studies. The extensive documentation of Cd's high toxicity and mobility in soil environments underscores its potential for causing significant ecological and human health hazards. For example, the work of (Wei, Jiang, Li, & Mu, 2010) emphasized the notable role of Cd in soil contamination, a significant factor causing ecological risk to the environment. Similarly, (V. Kumar, Sharma, et al., 2019) shed light on its bioaccumulative properties and serious health risks. The prominence of Cd in contributing to ecological risks in our study echoes these observations, further highlighting the adverse environmental impacts of this heavy metal. The significant roles of Pb and Cr in contributing to ecological risks are also supported by the existing literature (D.-M. Xu et al., 2017; Yarahmadi & Ansari, 2018). Lead, known for its persistence and capacity to inflict long-term harm on ecosystems, has been the subject of extensive research (Hu et al., 2019). (Jafari, Khorasani, & Danehkar, 2010), highlighted the cumulative nature of lead (Pb) and its far-reaching ecological consequences. This research underscores the importance of understanding the ecological impact of Pb contamination, as it can have lasting and widespread effects on the environment. Similarly, the role of Cr, particularly its hexavalent form, has been a topic of concern due to its carcinogenic properties and prevalence in industrial areas, as discussed by (J. Liu et al., 2009; Tian et al., 2020).



The SOM clustering analysis underlines the presence of unique heavy metal compositions in each cluster (I, II, III), suggesting that the ecological risks associated with these metals vary significantly across regions. Cluster I, located in the mid-reaches of the river within the study zone, is identified as the highest-risk area. The reasons for this heightened risk are multifaceted. The water quality of the Aik-stream in this area is severely compromised by industrial and urban waste. This pollution directly correlates with the soil's heavy metal ecological risk. Due to its location in a densely urbanized area with numerous industrial establishments, human activity significantly contributes to the ecological risk. It is consistent with the observations of (X.-Y. Zhou & Wang, 2019), who reported increased heavy metal pollution in urban and industrial regions due to anthropogenic activities. Moreover, the correlation between industrial and urban waste and soil heavy metal risk resonates with the findings of (Qing, Yutong, & Shenggao, 2015), highlighting the impact of urban runoff and industrial discharges.

Cluster I represents a primarily rural region with extensive agricultural activities. The key points include that, unlike Cluster II, human activity here is limited primarily to agriculture, resulting in a lower ecological risk from heavy metals. Though impactful, agricultural practices appear to pose less of a threat regarding heavy metal pollution than industrial activities. The lower ecological risk in the predominantly rural and agricultural Cluster I corroborates with the studies of (Y. Huang et al., 2018), which suggest that such areas generally experience lower heavy metal contamination. However, the contribution of agricultural practices to soil heavy metal levels, as pointed out by (Quinteros et al., 2017), indicates a need for careful management even in rural settings. Cluster III presents a moderate level of risk, influenced by a mix of environmental and human factors. Temperature and stream water quality align with the research of (Martínez- Alcalá & Bernal, 2020), who acknowledged the role of natural environmental factors alongside human activities in influencing heavy metal pollution. Cluster III showcases a different set of risk factors compared to Clusters I and II, indicating the complexity and variability of ecological risks in different environments.

Moreover, this study employs structural equation modeling (SEM) to investigate the interconnectedness of factors in soil, such as wastewater discharge, agricultural methods, and climatic conditions regarding metal accumulation. This methodology is in line with the approach used by (Hou et al., 2013). Recent research conducted by (Sharma & Raju, 2013) and (Jan, Ishaq, Ihsanullah, & Asim, 2010) have presented a connection between the presence



of metals in soil and the release of industrial wastewater. Our study reveals the impact of agricultural practices on the buildup of zinc (Zn) and copper (Cu) in soil, which is consistent with previous research conducted by (Liao et al., 2019) and (Genova et al., 2022). These researchers highlighted the impacts caused by these metals originating from agricultural activities. Moreover, our study provides a viewpoint in the field by shedding light on the impacts of climate conditions and soil quality on the concentrations of heavy metals. This perspective builds upon the work previously conducted. (X. Zhong et al., 2020), who explored the influence of environmental factors on pollutant dynamics. This research supports previous studies and provides additional insights, particularly in measuring how various environmental factors collectively play their role in the contribution of heavy metals in soil.

The study has provided insights and practical suggestions for research and policy-making in heavy metal contamination. To accurately assess the effectiveness of remediation efforts, conducting studies that track the changes in heavy metal concentrations over time is crucial. By broadening this study's scope and methodological approach to include regions worldwide, we can understand the ecological effects on a regional level. It is crucial to identify routes by which heavy metals enter the soil. This understanding is essential for developing strategies to mitigate their impact (Giuffré, Romaniuk, Marbán, Ríos, & Torres, 2012). The use of technologies, such as Geographic Information Systems (GIS) and machine learning, plays a role in improving the analysis and management of environmental challenges (Johnbull, Abbassi, & Zytner, 2019) and (H. Zhang et al., 2020).

Our research emphasizes the importance of employing targeted approaches in management and preservation when addressing the issue of metal pollution. One crucial aspect of this approach is examining plant life's role in environments. Understanding how various plants adapt to and mitigate soil pollution caused by metals is essential. This field of research is important for the resilience of plants in various environments and for creating environmentally friendly solutions to address the degradation of our surroundings.

### **3.5 Conclusion**

The study extensively assessed the presence of heavy metals in soil using advanced methods, like the Self Organizing Map (SOM) and Structural Equation Modeling (SEM). The main focus of this study was to examine the risks associated with eight heavy metals (Cu, Cd, Zn, Cr, Pb, Ni, As, and Hg) in the studied region. We utilized the Potential Ecological Risk



Index (PERI) as our measurement tool to achieve this. The findings showed serious risks associated with these heavy metals, particularly Cd, Cr, and Pb. These heavy metals accounted for 87.31% of the total ecological risk in the study area. Specifically, Cadmium (Cd) stood out as the concerning element, as indicated by its high PERI score of 258.5. Wastewater and agricultural activities were the factors that contributed to this risk, although there were also environmental factors that played indirect roles. This study emphasizes the necessity of tackling soil contamination caused by metals to mitigate its detrimental effects on the environment and ecosystems. Our integrated methodology combines SOM and SEM and provides a comprehensive approach to effectively assessing pollution levels across regions. This methodology allows us to understand soil contamination's extent and geographical pattern. The findings of this study are essential for policymakers and environmental managers because they offer a guide for making informed choices to protect soil health and tackle the issues associated with heavy metal pollution.





# Advanced Integrative Analytics for Data-Driven Risk Assessment of Ecological and Human Health Risks from Heavy Metal Contaminations in Soil

Ujala Ejaz<sup>1,2</sup>, Shujaul Mulk Khan<sup>1,3</sup> , Shah Fahad Ali Shah<sup>4</sup>, Noreen Khalid<sup>5,6</sup>,  
Sadia Jehangir<sup>1,2</sup>, Zarrin Fatima Rizvi<sup>5</sup>, Jens-Christian Svenning<sup>2</sup>

Show more

Add to Mendeley Share Cite

<https://doi.org/10.1016/j.hazadv.2025.100596>

Get rights and content

Under a Creative Commons license

open access

## Highlights

- Advanced Self-Organizing Maps (SOM) reveal significant spatial heterogeneity in soil contamination by heavy metals, necessitating targeted interventions.
- Cadmium is the main ecological risk factor, with lead and chromium also posing significant threats.



# **Understanding Vegetation Dynamics: Analyzing Structure, Composition, Distribution Patterns, Identifying Indicator Plant Species and the Role of Alien Species in Riparian Zone**

## **4.1 Introduction**

This chapter overviews the intricate dynamics of plant biodiversity in the environmentally challenged Aik-Stream region, characterized by heavy industrial wastewater pollution. This study examines the plant life thriving in the Aik-Stream's demanding environmental context, providing a comprehensive understanding of the environmental factors that shape biodiversity patterns within this pollution-affected landscape due to complex ecological interactions. Ecological research relies heavily on plant biodiversity, including a region's flora (O. H. Frankel, Brown, & Burdon, 1995). The rigorous floristic inventories conducted globally by researchers provide essential data for future botanical and ecological research. To ensure clarity in ecological discourse, it is imperative to clarify the distinction between the concepts of flora and vegetation. A flora is a systematic study of the diversity of plants within a region, usually the study of plants. Similarly, vegetation examines how these species are distributed and how abundance varies across geographical terrains (Khera, Kumar, Ram, & Tewari, 2001). It is essential to conduct floral surveys to understand various ecological aspects, including species diversity, soil characteristics, and climate conditions. Species diversity patterns and species distribution concerning environmental variables are analyzed by such surveys as a foundation of ecological research (Pérez-Escobar et al., 2017). Comprehensive vegetation classification, ecological services quantification, and mapping of vegetation patterns are essential outcomes of these studies (Reed, Sarasan, Kane, Bunn, & Pence, 2011). The study of plant biodiversity enhances our knowledge of the natural world and provides essential tools for conserving these vital resources (E. A. Ali, 1993). A holistic approach to biodiversity conservation is vital to guarantee its preservation for future generations in the event of extinction. Biological diversity represents a fundamental aspect of ecological conservation and sustainable development beyond academic inquiry (Chernen'kova, 2014).



#### ***4.1.1 Impact of Anthropogenic Activities on Natural Vegetation***

The impact of natural vegetation in geographic locales extends significantly beyond the influences of climate, geology, and soil typologies as it increasingly grapples with the pervasive effects of anthropogenic activities, notably environmental pollution (Chernen'kova, 2014). This pollution, arising from diverse sources such as industrial emissions, waste disposal practices, and the release of effluents, catalyzes profound and rapid transformations within natural ecosystems (Bayouli, Bayouli, Dell'Oca, Meers, & Sun, 2021). These transformations are characterized by marked changes in ecosystem structure, functionality, and species composition, thereby establishing pollution as a dominant and disruptive force within the natural environment. The detrimental impacts of pollution are particularly evident in the context of industrial activities, where the improper disposal of waste materials, especially those associated with mineral extraction processes, triggers a cascade of environmental repercussions. As elucidated by (S. Banerjee, Banerjee, Palit, & Roy, 2019), industrial pollution exerts a direct and far-reaching impact on natural vegetation, precipitating rapid alterations in the ecological balance. This scenario accentuates the critical need for developing, implementing, and rigorously enforcing environmental regulations and pollution control mechanisms, emphasizing international collaboration and compliance (Ahriz, Nedjraoui, & Sadki, 2010). The persistently growing impact of various pollutants places immense stress on natural ecosystems, representing a significant and complex challenge to environmental conservation and management (E. A. Ali, 1993). A particularly striking example of this impact is the degradation of forest ecosystems, which vividly illustrates the adverse effects of pollution on terrestrial biomes (Zvereva & Kozlov, 2011). This situation underscores the imperative for enhanced and concerted efforts in environmental stewardship to mitigate the harmful consequences of pollution on natural vegetation and ensure the long-term resilience and health of global ecosystems (Vorobeichik, 2022).

#### ***4.1.2 Indicator Species and Their Significance***

Without appropriate treatment and detoxification, effluents from the industries introduce potentially toxic elements, particularly heavy metals, into the ecosystem (Bunce et al., 2013). Over time, contamination levels in water and soil can escalate, adversely affecting the local natural environment and the well-being of local inhabitants. Pollution also profoundly impacts the composition of flora and fauna through chemical and physical alterations of the environment (Nally & Fleishman, 2004). Paradoxically, certain types of pollution can give rise to unique and diverse vegetation types in subtropical geographic



regions (Culmsee et al., 2014). Several plant species can absorb these toxic pollutants or potentially toxic elements from polluted ecosystems, serving as sinks that accumulate harmful concentrations of waste (W. Khan, Khan, Ahmad, Ahmad, & Page, 2016). These plants are termed "indicator species," "tolerant species," or "hyperaccumulators" of specific habitats (W. Khan et al., 2016).

In ecological research, a diverse range of plant species has been identified as crucial indicators for specific habitats, significantly aiding in determining optimal land use for varied applications such as forestry, agriculture, and mining (Cousins & Lindborg, 2004). This role of plant species as biological indicators in guiding land use practices is well-documented, with (S. M. Khan, Page, Ahmad, & Harper, 2014) providing key insights into their importance in making informed decisions regarding land utilization. These species are not merely of biological interest but are instrumental in bridging ecological research with practical land-use management, enhancing the sustainability of such practices (Hussain et al., 2019). Furthermore, plant indicators, as explored in the research by (Carignan & Villard, 2002), offer invaluable insights into the suitability of various soil types for specific agricultural applications. This research underscores the importance of understanding soil-plant interactions, vital for optimizing agricultural productivity and sustainability (Carignan & Villard, 2002). Complementing this, research into the impact of climatic factors on plant growth rates has revealed the potential of certain plants to serve as indicators for ecosystems or environments favorable to specific agricultural or forestry activities (Hussain et al., 2019). This aspect of research is critical in understanding the interplay between climatic conditions and plant physiology and its implications for land use planning (Nakamura et al., 2016).

Despite the recognized importance of plant indicators in ecological research, there has been a notable lack of a cohesive, comprehensive theoretical framework guiding their application. The concept of dominant or characteristic species within ecological communities has been mentioned in various studies but often lacks a systematic approach. This gap in methodology has been addressed by the Indicator Species Analysis (ISA), as proposed by (Dufrêne & Legendre, 1997). ISA provides a methodical and empirical approach for identifying and selecting plant indicators and assessing their performance across different groups of sampled units, emphasizing abundance and frequency. This method has significantly improved the precision and scientific robustness in selecting plant species as indicators, contributing substantially to ecological research and land use management (Xie, Zhang, Zeng, & He, 2020). The role of plant species as bioindicators is crucial in ecological



studies, providing essential guidance for sustainable land management practices (Güsewell, Peter, & Birrer, 2012). Researchers and land managers are equipped to make informed decisions by integrating botanical knowledge with ecological assessments, as facilitated by methodologies like ISA. These decisions aim to balance the needs of human development with the preservation of ecological integrity, thus contributing to the sustainable stewardship of natural resources (Chu et al., 2022).

In the wake of rapid industrial growth and escalating human population, the need to preserve our natural environment from overuse and pollution has become more pressing than ever. This urgency is particularly pronounced in regions like Sialkot, Pakistan, where the Aik-Stream has become a repository for industrial waste, profoundly influencing the local ecosystems. The contamination of the Aik-Stream has led to significant physical and chemical alterations in the water, adversely affecting both the natural ecosystems and the local communities reliant on these systems for their sustenance and wellbeing. The ramifications of such pollution are far-reaching, adversely affecting human health, wildlife, and plant life. The polluted waters pose many health risks, potentially leading to human and animal diseases and compromising food safety by contaminating water resources. Moreover, this pollution results in the degradation of crucial habitats. Therefore, it is imperative to undertake focused research to thoroughly evaluate the environmental impact of Aik-Stream's pollution, particularly on the local flora that thrives in these compromised conditions. Such research is crucial for developing effective strategies to alleviate these environmental challenges and ensure the health and longevity of the region's ecosystems and residents.

#### ***4.1.3 Anthropogenic Disturbances Drive Alien and Invasive Plant Dominance***

Understanding alien plant invasions in streamside vegetation remains incomplete for many regions, posing significant challenges for conservation efforts (Downey and Richardson 2016, Pathak et al. 2019). This unawareness hinders the recognition of plant community types more susceptible to invasion and the development of effective conservation strategies to preserve native vegetation (Reich et al. 2016). The complexity is compounded by the often poorly documented historical disturbance patterns and pristine ecological conditions of these riparian environments, making it difficult to assess changes induced by invasions (Downey and Richardson 2016). Moreover, it is often impractical to acquire consistent pre- and post-invasion data, either through longitudinal studies conducted at the same location or from comparable areas with similar vegetation but differing disturbance histories (Hejda et al. 2015). This makes direct contrasts between uninvaded and invaded streamside vegetation in



extensive landscapes rarely feasible (Diesburg 2021). The abundance and diversity of alien species in a specific area are shaped by numerous interacting factors, including the stage of invasion, the specific alien species present, the ecological characteristics of the community types, natural disturbance regimes, existing environmental conditions, and external disturbances (You et al. 2015). These complexities make it challenging to isolate specific mechanisms that promote invasions over large areas, which can impede immediate management responses. Despite these challenges, understanding the relationships between alien plant species and the attributes of streamside environments can reveal potential ecological processes and impacts of invasions. These insights can provide valuable indicators of where invasions are likely to be most severe, forming a crucial foundation for targeted management interventions in affected regions (Tonkin et al. 2018, Arif et al. 2021). This approach is adopted in this study to examine the role of alien and invasive species in the streamside vegetation along Aik-Stream, considering the topographic and climatic varied Chenab River basin in Punjab, Sialkot, Pakistan. In this region, descriptions of riparian vegetation from early or pre-settlement times are mostly anecdotal (Ullah et al. 2016, Ejaz et al. 2024b). Descriptions of plant community types exist for certain areas of the Sialkot region (Ikram et al. 2013) but the streamside vegetation along the Aik-Stream in Sialkot remains incompletely detailed. Moreover, the main part of currently established alien species in these plant communities is not well understood.

This research is centered on comprehensively examining the vegetation structure, composition, distribution patterns, and dynamics, identifying indicator species and the role of alien species in riparian zone impacted by anthropogenic activities in the Aik-Stream region of Sialkot, Pakistan. It addresses a notable gap in existing literature, which has primarily emphasized plant species as indicators of climatic conditions while considering the effects of industrial pollution on vegetation and their respective indicators. The hypothesis posits that the industrial wastewater-affected ecosystem in the Aik-Stream region harbors a distinct vegetation structure characterized by specific species that exhibit enhanced resilience, growth, and tolerance in polluted environments, as opposed to other plants in the same region. It is further theorized that the impact of industrial pollution in this area is significantly and positively correlated with these indicators and what is the contribution of alien and invasive species in the vegetation forms of streamside habitat? To delve into these hypotheses, the research focuses on the vegetation in the industrial wastewater-impacted ecosystem of the Aik-Stream region. A vital aspect of this study involves outlining the



statistical methodologies employed for identifying indicator plant species in this area, considering various environmental factors. The study utilizes Indicator Species Analysis (ISA) to identify indicator plants, supplemented by structural equation modeling and canonical correspondence analysis for the statistical substantiation of the hypothesized multivariate models. The methodology presented in this chapter is versatile and designed to be applicable for classifying and identifying indicator plants across diverse ecosystems globally. Our research objectives are multifaceted and focus on addressing key ecological and environmental challenges in the Aik-Stream region of Sialkot, Pakistan:

1. Conducting a comprehensive analysis of phytosociological attributes using advanced quantitative ecological methodologies to provide a detailed assessment of vegetation dynamics.
2. Evaluating the impacts of industrial wastewater on the structural composition, spatial distribution, and abundance of plant species, emphasizing anthropogenic influences on local biodiversity.
3. Identifying key indicator plant species across distinct vegetation zones by employing state-of-the-art ecological modelling techniques for accurate ecological characterization.
4. Investigating the intricate relationships between the richness and abundance of alien and invasive species and native plant diversity, while assessing their differential responses to environmental gradients and disturbance regimes.

## **4.2 Material and Methods**

### ***4.2.1 Vegetation sampling***

Quantitative ecological methods using quadrats were employed to sample the vegetation near the Aik stream area in Sialkot, Pakistan. The research was carried out around the Aik-Stream, which is affected by pollution produced by industries in Sialkot, Pakistan (Fig. 2.1). A total of 150 stations were randomly selected. Quadrat quantitative ecological techniques were implemented to sample vegetation. In total, 450 quadrats were sampled for vegetation purposes. At each station, quadrats measuring 100 meters, 25 meters, and 1 meter were collected to represent tree, shrub, and herb vegetation, respectively. Phytosociological characteristics such as plant cover, occurrence, density, relative occurrence, density, and importance value index were assessed for each plant species at every location. The cover and its relative values for tree species were calculated at the basal area of a stem through



Diameter at Breast Height techniques. Basal area was calculated using the formula  $\pi r^2$  (where r = radius) (Khan et al. 2014; Khan et al. 2013c; Khan et al. 2017b). All the reported plant species were collected, appropriately tagged, placed in a newspaper, and pressed in a plant presser (Ali and Nasir 1990; Ali and Qaiser 1995; Khan et al. 2013b; Khan et al. 2016). The Mercuric chloride and Ethyl alcohol solutions were used to poison specimens and mounted on standard herbarium sheets.

At last, all the plant specimens were identified with the help of Flora of Pakistan and other expert taxonomists (Nasir et al. 1972). Each subtropical station's geographical coordinates (longitude, latitude, and elevation) were recorded using GPS (Garmin etrex). A Geographical Information System (GIS) generated map was prepared for the sampling points using ArcGIS 10.5 software (Fu et al. 2013; Khan et al. 2016). The nativity of each taxon in Pakistan was determined using the Flora of Pakistan (Tropicos 2024). Plants were categorized as native (indigenous to Pakistan), alien (non-native to Pakistan), invasive (non-native and actively spreading), or cryptogenic species (those with uncertain origin or containing both native and non-native genetic variants in Pakistan) were grouped with non-native species for evaluations concerning non-native organisms. Plants with uncertain nativity, which usually occurred infrequently and had low cover, were clustered with native species. Species richness (S = number of species per plot), evenness (Pielou's J), and the Shannon diversity index (H') were computed for each plot (Hejda et al. 2009).

#### 4.2.2 Physiological attributes

Vegetation data, i.e., frequency, relative frequency, density, relative density, cover, relative cover and importance value index of each plant species in each quadrat, were measured by using the following formulas:

##### 4.2.2.1 Density

Density is the total number of individuals of a plant species in a sampled region. It was calculated according to (Khan 2012; Oosting 1956).

$$\text{Density (D)} = \frac{\text{Total no. of individuals of a species found in a quadrat}}{\text{Total sampled/quadrat area}} \dots \dots \dots (4.1)$$

##### 4.2.2.2 Relative density

Relative density is the % age distribution of individual species in a sampled area.

$$\text{Relative Density (RD)} = \frac{\text{Density of individuals plant species in a quadrat}}{\text{Total density of all plant species}} \times 100 \dots \dots \dots (4.2)$$



#### 4.2.2.3 Frequency

The percentage of number of a quadrats/sampled station in which a plant species is present is termed as frequency. It was calculated using the formula of (Cheevaporn and Menasveta 2003).

$$\text{Frequency (F)} = \frac{\text{No. of quadrats in which a plant species present}}{\text{Total no. of quadrats taken}} \times 100 \quad (4.3)$$

#### 4.2.2.4 Relative frequency

Relative frequency is the percentage frequency an individual species of the total frequencies.  $\text{Relative Frequency (RF)} = \frac{\text{Frequency of individual plant species}}{\text{Total frequency of all plant species}} \times 100$  (4.4)

#### 4.2.2.5 Cover

It is the basal area occupied by herbs, shrubs, or trees.

$$\text{Cover (C)} = \frac{\text{Total basal area of all individuals of a species}}{\text{Total quadrat/sampled area size}} \times 100 \quad (4.5)$$

#### 4.2.2.6 Relative cover

The percentage value obtained by dividing an individual's plant species' total cover by all plant species is termed relative cover.

$$\text{Relative Cover (RC)} = \frac{\text{Cover of individual plant species}}{\text{Total cover of all plant species}} \times 100 \quad (4.6)$$

#### 4.2.2.7 Basal area

The diameter of tree species was measured at Breast Height (in inches) using measuring tape at a height of 4.5 feet above the ground.

$$(cm) = CBH (inch) \times 2.54$$

According to the formula, meters require values to measure the basal area. The below conversion has been followed.

$$(m) = (cm) / 100$$

To find the radius, the value of 'r' was obtained by using the formula of circumference given below.

$$C = 2\pi r$$

This formula was rearranged as:



$$r = (m)/2\pi$$

The calculation of the basal area radius is required in square ( $r^2$ ). The derivation of the  $r^2$  is given as

$$r^2 = r \times r$$

All the derivations mentioned above were requisite for the formula for the basal area given below.  $Basalarea (BA) = \pi r^2$

#### 4.2.2.8 Cover classes for herb and shrub species

The percentage of cover for herbs and shrubs was estimated and noted in the field (Braun-Blanquet 1932). The percentage cover range and cover classes are below in Table 2.1, along with their midpoints.

Table 4.1 Braun Blanquet covers classes with their herb and shrub species midpoints.

Cover range (%)	Mid-point	Class
> 1%	0.5	Class $\infty$
1-5 %	3.0	Class 1
6-25 %	15.5	Class 2
26-50 %	38.0	Class 3
51-75 %	63.0	Class 4
75-100 %	88.0	Class 5

#### 4.2.3 Importance Value Index (IVI)

The IVI of each plant species at each station was calculated according to (Khan 2012). The relative density, frequency, and cover values were added and divided by 3. Mathematically, it is given as:

$$\text{Importance Value Index (IVI)} = RD + RF + RC / 3 \dots \dots \dots (4.7)$$

#### 4.2.4 Climatic Data

The climatic data, i.e., precipitation and temperature, were obtained from the Metrological Department of the Government of Pakistan.

#### 4.2.5 Soil Data Collection and Analysis

Soil samples were gathered from each station and taken at a depth of 0.3 meters (1 foot) using a soil sampling tool. The soil samples were then carefully placed in plastic bags



labeled and allowed to dry at room temperature. Afterward, an analysis was conducted to examine the chemical characteristics of the collected soil samples, such as Electrical Conductivity (EC) pH levels, Total Dissolved Solids (TDS) Organic matter content, Soil moisture levels, Potassium (K) concentration, Phosphorus (P) levels, Cadmium (Cd) presence, Nickle (Ni) composition, Chromium (Cr) content, Lead (Pb) amount, Copper (Cu) concentration, Zinc (Zn) quantity, Arsenic (As) levels and mercury (Hg). Soil EC, pH values, and TDS were determined using the McLean methods (McLean 1982). Ten grams of sifted and dried soil were mixed thoroughly in 50 milliliters of water using a magnetic stirrer for an hour. The mixture was then filtered using filter paper. The levels of conductivity (EC) pH and total dissolved solids (TDS) were measured using instruments: an EC meter (Adwa AD3000), a pH meter (Russel RL060P), and TDS meters. The content of matter was assessed through an analytical method (Tfaily et al. 2017). The concentration of the elements K, P, Ni, Cu, Cd, Pb, Cr, Zn, As and Hg was analyzed using atomic absorption spectroscopy (Ahmad et al. 2019). We placed a gram of dried sample into a 250 mL flask. Then, we added a mixture of 10 mL of Perchloric (HClO<sub>4</sub>) and Nitric (HNO<sub>3</sub>) acid in a 1:3 ratio. Let it sit for 24 hours. Soil samples were digested on a hot plate at an initial temperature of 150 °C for 1 hour and a final temperature of 235 °C until the red nitric acid fumes disappeared and white fumes appeared. The solution was filtered after cooling through filtered paper (Whatman No. 42), and 40 mL of distilled water was added to raise its volume. The blank reagents were also prepared. The atomic absorption spectrophotometer VARIAN, AA240FS, was used for the elemental analyses. The final element concentration was obtained using the formula below:

$$\text{Element concentration (mg/kg)} = \frac{\text{AA reading} - \text{Blank reagents}}{\text{sample weight (kg)} \times \text{volume raised} \times \text{df}} \dots\dots\dots (4.8)$$

Where AA= atomic absorption reading, df = dilution factor.

#### **4.2.6 Statistical Analysis**

All the collected plant species, stations and environmental data of the Aik-Stream region were analyzed using different multivariate statistic software, i.e., PC-ORD v.5, Canoco v 4.5, SPSS v.20, STATA v.14, and R v.4.0.2. Plant data along with environmental variables were arranged in Microsoft EXCEL worksheet for further analyses. The detailed description of the subsequent analyses is provided below:



#### 4.2.6.1 Species-area curves

PC-ORD version 5 was used to draw species-area curves. This was performed to establish whether the sample sizes were adequate or not. Species area curves are primarily utilized in vegetation ecology to realize species composition in relation to sample size. Plant abundance data with Sorensen distance values were used to create species-area curves of the studied area.

#### 4.2.6.2 Cluster Analysis

The cluster analysis (CA) used the total vegetation data with their environmental variable to classify the whole data into respective vegetation zones, and two-way cluster analysis was used to identify significant vegetation zones and distribution of plant species at individual levels in each quadrat. This analysis was done using PCORD V5. This was based on the pattern similarity index through the Sorensen distance measurement and Wards Linkage Method (Ahmad et al. 2016b; Greig-Smith 1983; Khan et al. 2016).

#### 4.2.6.3 Indicator Species Analysis (ISA)

The ISA was carried out to find indicators of each vegetation zone of the Aik-Stream. It provides knowledge about species fidelity with the particular habitat of a specific vegetation zone. The Monte Carlo test was used for statistical significance after determining indicator values (%age of perfect indication established on combining values of relative abundance and frequency) of respective indicators using a method initially adopted in a study (Dufrene and Legendre 1997). During ISA, a proportional abundance of a specific plant in a particular group, its relative abundance was calculated using the formula given below:

$$RA_{jk} = \frac{x_{kj}}{\sum_{k=1}^g x_{kj}} \dots \dots \dots (4.9)$$

Where  $RA_{jk}$ =relative abundance,  $X_{kj}$ = means an abundance of species j in group k, g=total number of groups.

Then, the relative frequency of plants in each group was also calculated, i.e., the proportion of sample units in each group containing that plant species using the formula below. Percent/faithfulness/constancy of presence in a particular group is also expressed using these procedures:

$$RF_{kj} = \frac{\sum_{i=1}^{n_k} b_{ijk}}{n_k} \dots \dots \dots (4.10)$$



Where  $RF_{kj}$  is the relative frequency of plant  $j$  in group  $k$ ,  $b_{ijk}$  is the presence or absence of plant  $j$  in sample  $i$  of group  $k$ ,  $i$  is the sample unit.

At last, equations ix and x were gathered multiplication, and the results were expressed as percentage yielding indicator value ( $IV_{kj}$ ) for each plant  $j$  in group  $k$ .

$$IV_{kj} = 100(RA_{jk} \times RF_{kj}) \dots \dots \dots (4.11)$$

A threshold level of 25% indication and 95% significance ( $p \leq 0.05$ ) was deliberated as a cutoff value for determining the indicators. Once the significant indicators were identified, the direct gradient analysis i.e., CCA, was performed using CANOCO software to examine, reconfirm and draw the substantial and distinct indicators of each sort of vegetation zone of the Aik-Stream. CCA analyzes the indicator plant's relation by multiple linear regression with environmental gradients and gives us an interpretable graphical presentation of species response environmental variables (Dufrêne and Legendre 1997; Ter Braak and Prentice 1988).

#### 4.2.6.4 Diversity Indices, Species richness and evenness

Diversity indices provide a mathematical estimation of plant species diversity in a community. This gives us more information concerning community composition. The diversity indices i.e., Shannon Index ( $H'$ ), Simpson Index ( $D$ ), Simpson Index of diversity ( $1-D$ ) and Pielou's Evenness ( $J$ ) were calculated (Pielou 1966; Simpson 1949).

##### 4.2.6.4.1 Shannon Index ( $H'$ )

The Shannon index is usually used to know about plant species diversity in a particular community. It determines the evenness and abundance of plant species in a community. The Shannon index of diversity was calculated using the following:

$$H' = \sum_{i=1}^S p_i \ln p_i \dots \dots \dots (4.12)$$

Where  $H'$ = Shannon Index,  $S$ = Total number of species in sample/community/zone (species richness),  $p_i$ = Relative abundance of each species.

##### 4.2.6.4.2 Simpson Index ( $D$ )

The Simpson index also measures plant species diversity. It was calculated using the below equation:

$$D = \frac{\sum n_i (n_i - 1)}{N(N - 1)} \dots \dots \dots (4.13)$$



Where D=Simpson index, n= total number of any particular plant species, N= Total number of all species.

#### 4.2.6.4.3 Simpson Index of Diversity (1-D)

The Simpson index of diversity was determined using 1-D (D is the Simpson index). Its value ranges between 0-1.

#### 4.2.6.4.4 Pielou's Evenness (J) / Species Evenness

Pielou's evenness index was used to determine plant species' evenness using the equation below.

$$J = H' / H_{max} \dots \dots \dots (4.14)$$

Where J= Pielou evenness, H'= Calculated Shannon Index, Hmax= ln(s) [species diversity under maximum equitability].

#### 4.2.6.4.5 Species richness

The species richness was calculated using the below equation (Menhinick, 1964).

$$d = \frac{s}{\sqrt{n}} \dots \dots \dots (4.15)$$

Where s=total number of species, N=total number of individuals, and d=species richness.

#### 4.2.6.4.6 Brillouin Diversity Index

The Brillouin Diversity Index, often used in ecological contexts, quantifies species diversity within a community. It considers the number of species (S) and their evenness in abundance (ni). The formula for calculating the Brillouin Diversity Index (B) I is as follows:

$$B = \ln(S) - \frac{1}{2N} \sum_{i=1}^s (n_i - 1) \ln(n_i) \dots \dots \dots (4.16)$$

Whereas B indicates that the Brillouin Diversity Index, ln is the natural log, S is the number of species, N is the number of individuals, and ni is the total number of abundances of the ith species.

#### 4.2.6.4.7 Margalef Diversity Index

The Margalef Diversity Index is another measure of species' diversity in ecological communities. This index measures species richness and considers the total number of species



(S) and total number of individuals (N) in a community. The formula for this index (D<sub>M</sub>) is as follows:

$$D_M = \frac{(S - 1)}{\ln(N)} \dots \dots \dots (4.17)$$

The eq. D<sub>M</sub> is the Magalef index, the total number of species denotes S, and N is the total number of individuals.

#### 4.2.6.4.8 *Equitability-J Index*

The equitability-j diversity index measures the diversity that incorporates species richness and evenness in a community. The formula is as follows:

$$J = \frac{H'}{\ln(S)} \dots \dots \dots (4.18)$$

Where J indicates the equitability diversity index, H is the Shannon-Weaver diversity Index and S is the number of species.

#### 4.2.6.4.9 *Fisher's Alpha (α)*

The Fisher's Alpha measures alpha diversity, quantifying species diversity within the specific community.

$$\alpha = \ln(S) - \frac{N_1}{N} \dots \dots \dots (4.19)$$

The eq (α) in the Fisher diversity and ln the natural log of the total numbers of species in the study area.

#### 4.2.6.4.10 *Berger-Parker Index*

The Berger-Pakker index represents the dominance of the most abundant species in a community. The formula is:

$$d = \frac{N_{max}}{N} \dots \dots \dots (4.20)$$

Whereas the d is the Berger-Parker index  $N_{max}$  is the abundance of the most abundant species, and N is the total number of individuals.

#### 4.2.6.4.11 *Chao-1 Estimator*

Chao-1 estimates the species richness based on the number of singletons and doubletons. The formula is as follows:



$$Chao - 1 = S_{obs} + \frac{F_1^2}{2F_2} \dots \dots \dots (4.21)$$

Chao-1 is the estimated species richness; S<sub>obs</sub> is the total observed species, F<sub>1</sub> is the number of singletons and F<sub>2</sub> is the number of doubletons.

#### 4.2.6.4.12 iChao-1 Estimator

The iChao-1 is the advanced version of the Chao-1 estimator. The formula is a variation of Chao-1:

$$iChao - 1 = S_{obs} + \frac{F_1(F_1 - 1)}{F_2(F_2 + 1)} \dots \dots \dots (4.22)$$

#### Abundance-based Coverage Estimator (ACE)

The ACE estimates the total number of species based on the abundance distribution; the formula is:

$$ACE = S_{obs} + \frac{F_1^2}{2(F_2 + F_1)} \dots \dots \dots (4.23)$$

ACE represents the estimated species richness, S<sub>obs</sub> is the total observed species, F<sub>1</sub> is the number of singletons and F<sub>2</sub> is the number of doubletons.

### 4.3 Results

A total of 182 plant species belonging to 136 genera and 49 different plant families were recorded around Aik-Stream of Sialkot, Pakistan. The detailed description and zone-wise composition and distribution pattern are discussed stepwise below:

#### 4.3.1 Plant species composition

Habit wise, the recorded plant species comprised 127 herbs (70% of the total vegetation), 26 shrubs (14%) and 28 trees (16%) (Fig. 4.1; Table 4.1 and Appendix Table 1). The most dominant family in our study was Poaceae, represented by 19 species (13%). The second most dominant family was Compositae, represented by 18 members (13%), while Fabaceae has 16 species. The other families had a variable number of species, such as eight species of the Moraceae, while seven members represented Araceae. The least dominant families are Acanthaceae, Anacardaceae, Aopocynaceae, Cannabaceae, etc., each represented by one species. They shared less than 2 % of the total vegetation around Aik-Stream.



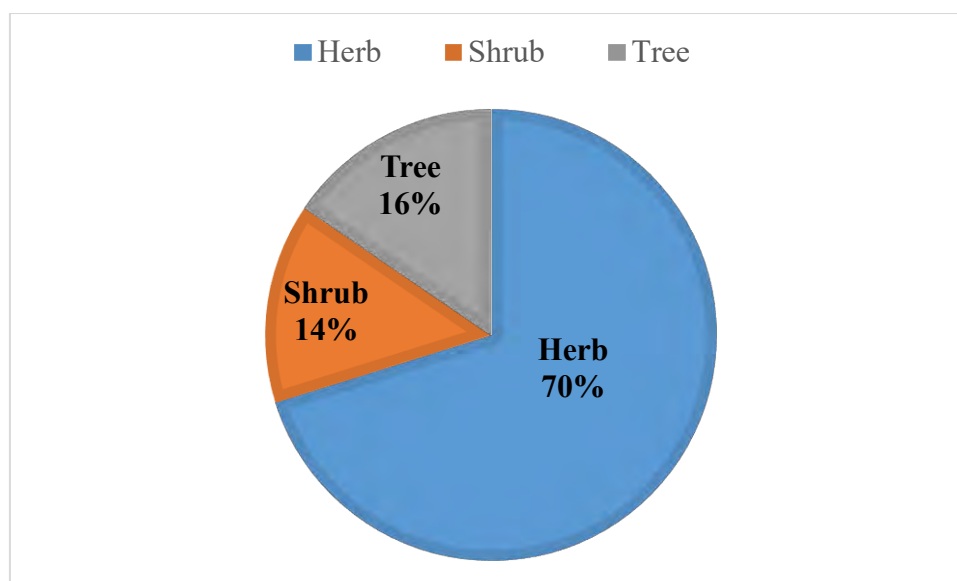


Figure 4.1 Percentage distribution of herbs, shrubs and trees around Aik-Stream of Sialkot, Pakistan.

Table 4.2 Detail of plant species and their habit family reported around Aik-Stream of the Sialkot, Pakistan.

S. No.	Plant Species	Habit	Family
1	<i>Abelmoschus moschatus</i> Medic.	Herb	Malvaceae
2	<i>Abutilon indicum</i> (L.) Sweet	Herb	Malvaceae
3	<i>Acacia homalophylla</i> Medic.	Tree	Fabaceae
4	<i>Acacia nilotica</i> (L.) Delile	Tree	Fabaceae
5	<i>Achyranthes aspera</i> L.	Herb	Amaranthaceae
6	<i>Adiantum capillus-veneris</i> L.	Herb	Pteridaceae
7	<i>Ageratum conyzoides</i> L.	Herb	Compositae
8	<i>Alocasia macrorrhizos</i> (L.) G.Don	Herb	Araceae
9	<i>Amaranthus graecizans</i> L.	Herb	Amaranthaceae
10	<i>Amaranthus retroflexus</i> L.	Herb	Amaranthaceae
11	<i>Amaranthus spinosus</i> L.	Herb	Amaranthaceae
12	<i>Amaranthus viridis</i> L.	Herb	Amaranthaceae
13	<i>Anagallis arvensis</i> L.	Herb	Primulaceae
14	<i>Argemone mexicana</i> L.	Herb	Papaveraceae
15	<i>Artemisia brevifolia</i> L.	Herb	Compositae
16	<i>Artemisia scoparia</i> Waldst. & Kitam	Herb	Compositae



17	<i>Arundo donax</i> L.	Shrub	Poaceae
18	<i>Avena sativa</i> L.	Herb	Poaceae
19	<i>Bergia capensis</i> L.	Herb	Elatinaceae
20	<i>Bidens bipinnata</i> L.	Herb	Compositae
21	<i>Bidens pilosa</i> L.	Herb	Compositae
22	<i>Boerhavia procumbens</i> Banks ex Roxb.	Herb	Nyctaginaceae
23	<i>Bougainvillea spectabilis</i> Willd.	Shrub	Nyctaginaceae
24	<i>Brachiaria reptans</i> L.	Herb	Poaceae
25	<i>Brachiaria ramosa</i> L.	Herb	Poaceae
26	<i>Brassica compestris</i> L.	Herb	Brassicaceae
27	<i>Brassica oleracea</i> L.	Herb	Brassicaceae
28	<i>Bromus japonicus</i> Thunb.	Tree	Poaceae
29	<i>Brossunatia papyrifera</i> (L.) L'Hér. ex Vent.	Tree	Moraceae
30	<i>Callistemon acerifolius</i> Mouill.	Shrub	Myrtaceae
31	<i>Calotropis procera</i> (Aiton) Dryand.	Shrub	Apocynaceae
32	<i>Campylanthus ramosissimus</i> Wight	Herb	Plantaginaceae
33	<i>Cannabis sativa</i> L.	Shrub	Cannabaceae
34	<i>Capsella bursa-pastoris</i> Medik.	Herb	Brassicaceae
35	<i>Cassia occidentalis</i> L.	Herb	Fabaceae
36	<i>Catharanthus roseus</i> (L.) G.Don	Shrub	Apocynaceae
37	<i>Cenchrus biflorus</i> Roxb.	Herb	Poaceae
38	<i>Cenchrus ciliaris</i> L.	Herb	Poaceae
39	<i>Cestrum nocturnum</i> L.	Shrub	Solanaceae
40	<i>Chenopodium album</i> L.	Herb	Amaranthaceae
41	<i>Cichorium intybus</i> L.	Herb	Compositae
42	<i>Clinopodium umbrosum</i> (M.Bieb.) Kuntze.	Herb	Lamiaceae
43	<i>Colocasia esculenta</i> (L.) schott	Herb	Araceae
44	<i>Colocasia gigantea</i> (Blume) Hook.f.	Herb	Araceae
45	<i>Combretum indicum</i> (L.) DeFilipps	Shrub	Combretaceae



46	<i>Commelina benghalensis</i> L.	Herb	Commelinaceae
47	<i>Convolvulus arvensis</i> L.	Herb	Convolvulaceae
48	<i>Conyza bonariensis</i> (L.) Cronquist	Herb	Compositae
49	<i>Conyza canadensis</i> (L.) Cronq.	Herb	Compositae
50	<i>Coronops didymus</i> (L.) Sm.	Herb	Brassicaceae
51	<i>Cynodon dactylon</i> (L.) Pers.	Herb	Poaceae
52	<i>Cynodon radiatus</i> Roth.	Herb	Poaceae
53	<i>Cyperus rotundus</i> L.	Herb	Cyperaceae
54	<i>Dalbergia sisso</i> DC	Tree	Fabaceae
55	<i>Daphane macronata</i> Royle	Herb	Thymelaeaceae
56	<i>Datura alba</i> L.	Herb	Solanaceae
57	<i>Datura metel</i> Moc. & Sessé ex Dunal	Herb	Solanaceae
58	<i>Datura innoxia</i> Mill.	Shrub	Solanaceae
59	<i>Datura metel</i> L.	Shrub	Solanaceae
60	<i>Debregeasia saeneb</i> (Forssk.)	Shrub	Urticaceae
61	<i>Desmostachy bippinanta</i> (L.) Stapf	Herb	Poaceae
62	<i>Dicanthium annulatum</i> (Forssk) Stapf	Herb	Poaceae
63	<i>Duranta stenostachya</i> Tod	Shrub	Verbenaceae
64	<i>Duranta erecta</i> L.	Shrub	Verbenaceae
65	<i>Dysphania ambrosioides</i> (L.)	Herb	Amaranthaceae
66	<i>Echeandia reflexa</i> (Cav.) Rose	Herb	Asparagaceae
67	<i>Echinops latifolius</i> Tausch	Herb	Compositae
68	<i>Eclipta alba</i> (L.) Hassk	Herb	Compositae
69	<i>Eichhornia crassipes</i> (Mart.) Solms-Laub	Herb	Pontederiaceae
70	<i>Erigeron canadensis</i> L.	Herb	Compositae
71	<i>Erigeron bonariensis</i> L.	Herb	Compositae
72	<i>Erythrina crista-galli</i> L.	Tree	Fabaceae
73	<i>Eucalyptus camaldulensis</i> Labill.	Tree	Myrtaceae
74	<i>Eucalyptus globulus</i> Labill.	Tree	Myrtaceae
75	<i>Euphorbia helioscopia</i> L.	Herb	Euphorbiaceae
76	<i>Euphorbia hirta</i> L.	Herb	Euphorbiaceae



77	<i>Ficus benghalensis</i> L.	Tree	Moraceae
78	<i>Ficus carica</i> L.	Tree	Moraceae
79	<i>Ficus elastica</i> Roxb.	Tree	Moraceae
80	<i>Ficus religiosa</i> L.	Tree	Moraceae
81	<i>Ficus virens</i> Aiton.	Tree	Moraceae
82	<i>Ficus macrophylla</i> L.	Tree	Moraceae
83	<i>Ficus palmata</i> Forssk	Tree	Moraceae
84	<i>Fumaria indica</i> (Hausskn.) Pugsley	Herb	Papaveraceae
85	<i>Galium aparine</i> L.	Herb	Rubiaceae
86	<i>Geum urbanum</i> L.	Herb	Rosaceae
87	<i>Goldbachia laevigata</i> (M. Bieb.) DC.	Herb	Brassicaceae
88	<i>Heracleum sphondylium</i> L.	Herb	Apiaceae
89	<i>Imperata cylindrical</i> (L.) Raeusch.	Herb	Poaceae
90	<i>Indigofera linifolia</i> (L.f.) Retz.	Herb	Fabaceae
91	<i>Ipomia purpurea</i> (L.) Roth	Herb	Convolvulaceae
92	<i>Ipomoea carnea</i> Jacq.	Herb	Convolvulaceae
93	<i>Jasminum nudiflorum</i> L.	Shrub	Oleaceae
94	<i>Juncus acuminatus</i> L.	Herb	Juncaceae
95	<i>Juncus effuses</i> L.	Herb	Juncaceae
96	<i>Jurinea heteromalla</i> (D.Don) N.Garcia, Herrando	Herb	Compositae
97	<i>Justica nilgherrensis</i> L.	Herb	Acanthaceae
98	<i>Koelaria macarantha</i> (Ledeb.) Schult.	Herb	Poaceae
99	<i>Lantana camara</i> L.	Shrub	Verbenaceae
100	<i>Camara vulgaris</i> Benth.	Shrub	Verbenaceae
101	<i>Lathyrus aphaca</i> L.	Herb	Fabaceae
102	<i>Lathyrus pseudocicera</i> Pomp.	Herb	Fabaceae
103	<i>Lemna minor</i> L.	Herb	Araceae
104	<i>Leuccocasia gigantea</i> (Blume) Schott	Herb	Araceae
105	<i>Malava neglecta</i> Wallr.	Herb	Malvaceae
106	<i>Malvastrum coromandelianum</i> (L.) Garcke	Herb	Malvaceae



107	<i>Mangifera indica</i> L.	Tree	Anacardiaceae
108	<i>Marsilea mutica</i> Mett.	Herb	Marsileaceae
109	<i>Medicago denticulata</i> Willd	Herb	Fabaceae
110	<i>Medicago minima</i> (L.) Grub.	Herb	Fabaceae
111	<i>Medicago polymorpha</i> L.	Herb	Fabaceae
112	<i>Melia azedarach</i> L.	Tree	Meliaceae
113	<i>Mentha spicata</i> L.	Herb	Lamiaceae
114	<i>Morus alba</i> L.	Tree	Moraceae
115	<i>Morus nigra</i> L.	Tree	Moraceae
116	<i>Nandina domestica</i> L.	Shrub	Berberidaceae
117	<i>Nastrum officinales</i> R. Br.	Herb	Brassicaceae
118	<i>Nerium oleander</i> L.	shrub	Apocynaceae
119	<i>Oxalis corniculata</i> L.	Herb	Oxalidaceae
120	<i>Parthenium hysterophorus</i> L.	Herb	Compositae
121	<i>Parthenocissus inserta</i> L.	Shrub	Vitaceae
122	<i>Paspalum paspalodes</i> (Michx.)	Herb	Poaceae
123	<i>Persiaria glabra</i> L.	Herb	Polygonaceae
124	<i>Persicaria hydropiper</i> L.	Herb	Polygonaceae
125	<i>Persicaria maculosa</i> Gray	Herb	Polygonaceae
126	<i>Phalaris minor</i> Retz.	Herb	Poaceae
127	<i>Phragmites karka</i> (Retz.) Trin. ex Steud.	Herb	Poaceae
128	<i>Phyla nodiflora</i> (L.) Greene	Herb	Verbenaceae
129	<i>Physalis divaricata</i> D. Don	Herb	Solanaceae
130	<i>Pistia stratiotes</i> L.	Herb	Araceae
131	<i>Plantago minor</i> L.	Herb	Plantaginaceae
132	<i>Populus alba</i> L.	Tree	Salicaceae
133	<i>Populus nigra</i> L.	Tree	Salicaceae
134	<i>Prosopis juliflora</i> (Sw.) DC	Tree	Fabaceae
135	<i>Ranunculus muricatus</i> L.	Herb	Ranunculaceae
136	<i>Riccinus communis</i> L.	Shrub	Euphorbiaceae
137	<i>Rifolium microdon</i> Hook. & Arn.	Herb	Fabaceae



138	<i>Rosa indica</i> L.	Shrub	Rosaceae
139	<i>Rosa webbiana</i> L.	Shrub	Rosaceae
140	<i>Rubus fruticosus</i> L.	Shrub	Rosaceae
141	<i>Rumex dentatus</i> L.	Herb	Polygonaceae
142	<i>Rumex hastatus</i> L.	Shrub	Polygonaceae
143	<i>Rumex nepalensis</i> Spreng.	Herb	Polygonaceae
144	<i>Saccharum bengalensis</i> Retz	Shrub	Poaceae
145	<i>Saccharum spontaneum</i> L.	Shrub	Poaceae
146	<i>Salix tetrasperma</i> Roxb.	Tree	Salicaceae
147	<i>Salsola Kali</i> L.	Herb	Amaranthaceae
148	<i>Senna occidentalis</i> (L.) Link	Herb	Fabaceae
149	<i>Setaria pumila</i> (Poir.) Roem. & Schult	Herb	Poaceae
150	<i>Sida cordata</i> (Burm.f.) Borss. Waalk	Herb	Malvaceae
151	<i>Silybum marianum</i> (L.) Gaertn.	Herb	Compositae
152	<i>Sinapis arvensis</i> L.	Herb	Brassicaceae
153	<i>Sisimbrium irio</i> L.	Herb	Brassicaceae
154	<i>Solanum lycopersicum</i> L.	Herb	Solanaceae
155	<i>Solanum nigrum</i> L.	Herb	Solanaceae
156	<i>Sonchus asper</i> L.	Herb	Compositae
157	<i>Sonchus oleraceous</i> L.	Herb	Compositae
158	<i>Sorghum halepense</i> (L.) Pers	Herb	Poaceae
159	<i>Stellaria media</i> L.	Herb	Caryophyllaceae
160	<i>Sylibum marianum</i> (L.) Gaertn.	Herb	Compositae
161	<i>Syzygium cumini</i> (L.) Skeels	Tree	Myrtaceae
162	<i>Tamarix dioica</i> Roxb. Ex Roth	Tree	Tamaricaceae
163	<i>Taraxacum officinale</i> L.	Herb	Compositae
164	<i>Torilis japonica</i> (Houtt.) DC	Herb	Apiaceae
165	<i>Torilis leptophylla</i> (L.) Rchb.f	Herb	Apiaceae
166	<i>Trifolium alexandrinum</i> L.	Herb	Fabaceae
167	<i>Trifolium resupinatum</i> L.	Herb	Apiaceae
168	<i>Triticum aestivum</i> L.	Herb	Poaceae
169	<i>Typha angustifolia</i> L.	Herb	Typhaceae



170	<i>Verb supina</i> L.	Herb	Verbenaceae
171	<i>Verbascum songaricum</i> Schrenk	Herb	Scrophulariaceae
172	<i>Verbascum thapsus</i> L.	Herb	Scrophulariaceae
173	<i>Verbena bonariensis</i> L.	Herb	Scrophulariaceae
174	<i>Verbena officinale</i> L.	Herb	Verbenaceae
175	<i>Veronica anagallis-aquatica</i> L.	Herb	Plantaginaceae
176	<i>Vicia sativa</i> L.	Herb	Fabaceae
176	<i>Vitis vinifera</i> L.	Shrub	Vitaceae
177	<i>Withania somnifera</i> (L.) Dunal	Herb	Solanaceae
178	<i>Withania somnifera</i> L.	Shrub	Solanaceae
179	<i>Xanthium strumarium</i> L.	Herb	Compositae
180	<i>Ziziphus jujuba</i> Mill.	Tree	Rhamnaceae
181	<i>Ziziphus jujuba</i> Mill.	Tree	Rhamnaceae
182	<i>Ziziphus nummularia</i> (Burm.f.) Wight & Arn.	Tree	Rhamnaceae

#### 4.3.2 Species area curves

Plant species classification into potential plant zones using PC-ORD version 5 involved several key steps. The initial examination involved a species-area curve to assess the adequacy of the sample size, a technique commonly employed in vegetation ecology for understanding species composition based on sample size. This analysis utilized plant abundance data and Sorensen distance values, encompassing 150 quadrats/stations and 182 species. The first-order jack-knife estimate was evident at station 135, followed by the second-order jack-knife estimate at station 145. Among the 182 species, 54 were observed to have a single occurrence in the study region. Notably, the maximum number of plant species was observed up to station number 135, totaling 170 species within an average distance of 0.039. Moving forward to station number 145, the average number of plant species recorded was 178, with an average distance of 0.024. Beyond this point, the species curves displayed parallel trends, with only four additional species recorded. These findings strongly indicated that the sampling efforts in the targeted region were sufficient, as illustrated in Figure 4.2.



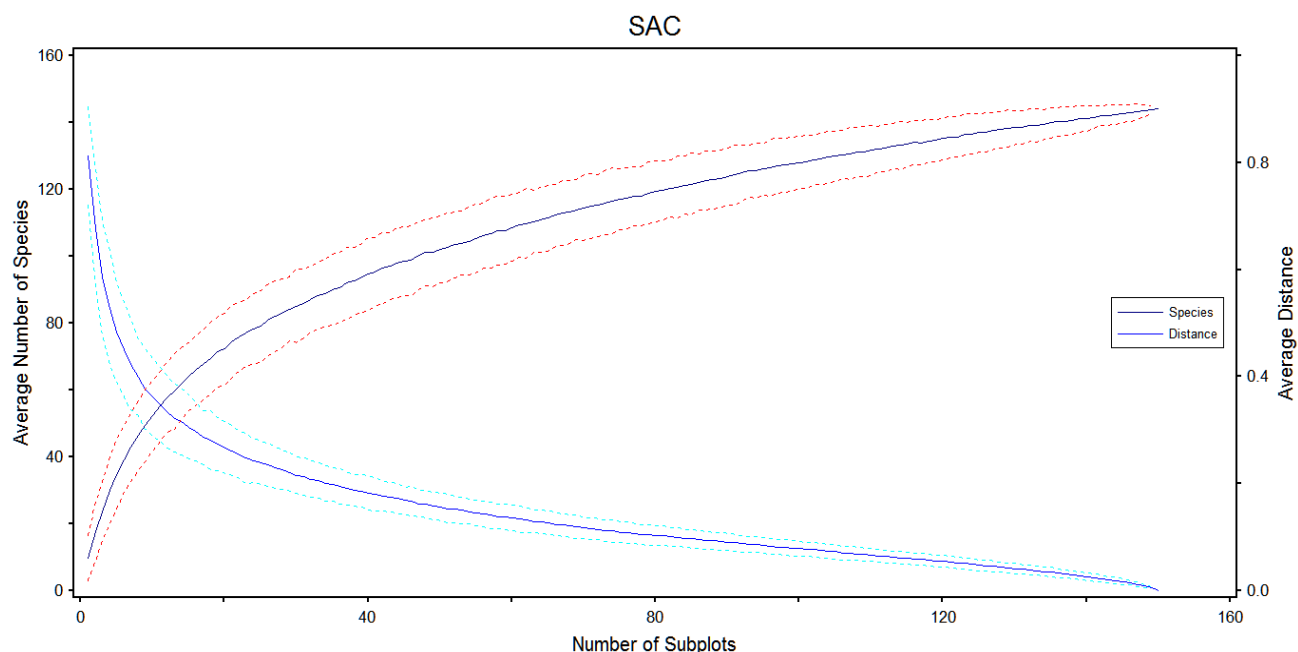


Figure 4.2 The species-area curves of 182 plant species distributed among 150 stations in the Aik-Stream region of Sialkot, Pakistan.

#### 4.3.3 Abundant and rare plants of the Aik-Stream

The abundant and rare plant species were identified based on the importance value index. Their detailed descriptions are as follows:

##### 4.3.3.1 Abundant and rare trees

In the Aik-Stream region of Sialkot, Pakistan, a comprehensive assessment of tree species was conducted, utilizing the Importance Value Index (IVI) as a key metric. Among these species, a select group emerged as the dominant and abundant, as indicated by their notably higher IVI scores. Topping the list was *Dalbergia sisso*, with an IVI of 1928.2, followed closely by *Acacia nilotica* at 711.5 IVI. The list of dominant species also included *Ziziphus Mauritania* (IVI: 679.88), *Brossunatia papyrifera* (IVI: 666.4), *Ficus carica* (IVI: 560.0), *Morus nigra* (IVI: 557), *Morus alba* (IVI: 434.15), *Acacia homalophylla* (IVI: 224.6), and *Ficus benghalensis* (IVI: 213.8), as depicted in Figure 4.3. Conversely, the study identified the top rarest tree species within the same region, characterized by their lower IVI values. *Eucalyptus globulus* emerged as the rarest, with an IVI of 42.74, followed by *Syzygium cumini* (IVI: 40.7) and *Mangifera indica* (IVI: 30.62). Other less common species included *Ficus macrophylla* (IVI: 21.54), *Salix tetrasperma* (IVI: 21.38), *Bromus japonicus* (IVI: 20.6), *Ficus virens* (IVI: 18.87), *Ziziphus jujuba* (IVI: 8.75), *Melia azedarach* (IVI: 18.40), and *Ficus palmata* (IVI: 16.43). Notably, *Populus nigra*, with an IVI of 11.46, also



ranked among the rarer tree species in this region. This comprehensive analysis is visually represented in Figure 4.4, shedding light on tree species' relative abundance and scarcity in the Aik-Stream region of Sialkot, Pakistan.

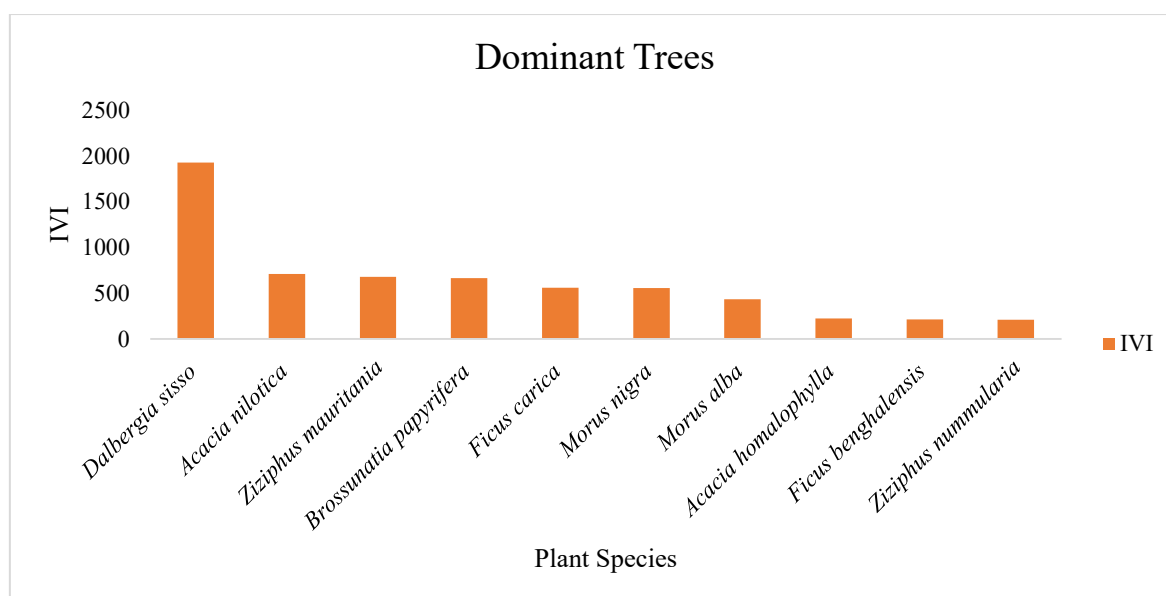


Figure 4.3 The topmost abundant tree species around the Aik-stream, Sialkot, Pakistan.

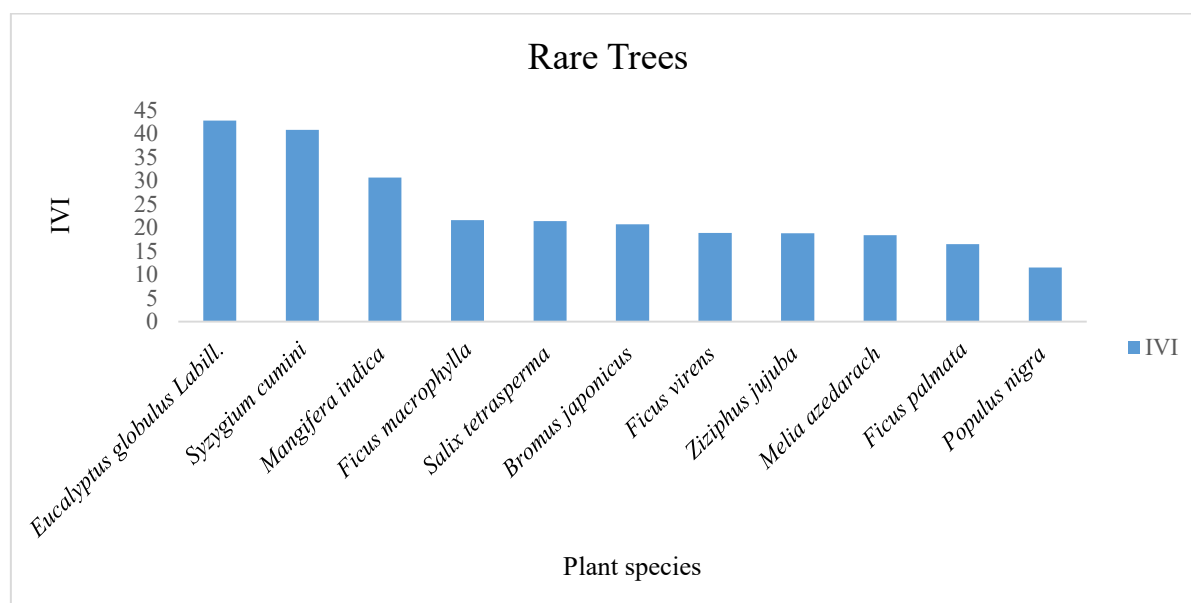


Figure 4.4 The rare tree species around the Aik stream, Sialkot, Pakistan.

#### 4.3.3.2 Abundant and rare shrubs

The most abundant shrub species included *Cannabis sativa L.* (with an IVI of 3340), *Arundo donax* (IVI of 1349.58), *Riccinus communis* (IVI of 1308.3), *Saccharum bengalensis* (IVI of 1107.16), *Parthenocissus inserta* (IVI of 386.40), *Lantana camara* (IVI of 344.39), and *Calotropis procera* (IVI of 325.12) in the Aik-Stream region (Fig. 3.4), Conversely, rare



shrub species included *Duranta stenostachya* (IVI of 19.598), *Debregeasia saeneb* (IVI of 13.5743), *Datura metel* (IVI of 13.3), *Nandina domestica* (IVI of 12.75), *Rubus fruticosus* (IVI of 12.753), *Jasminum nudiflorum* (IVI of 11.54), *Rumex hastatus* (IVI of 10.838), and *Nerium oleander* (IVI of 8.534) (Fig. 4.5)

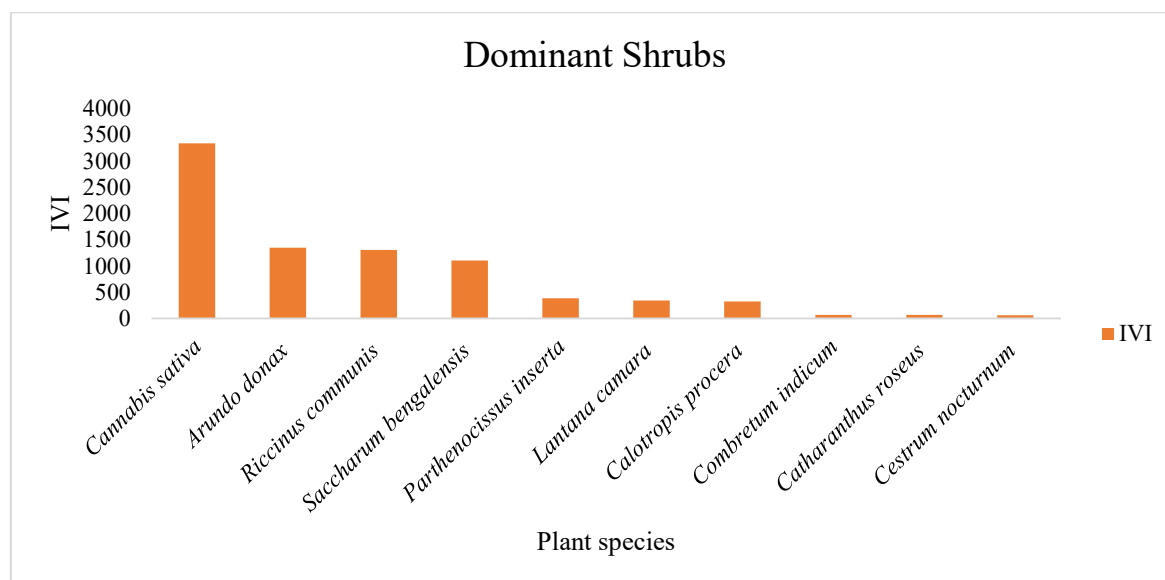


Figure 4.5 The Dominant shrubs species around the Aik-stream, Sialkot, Pakistan.

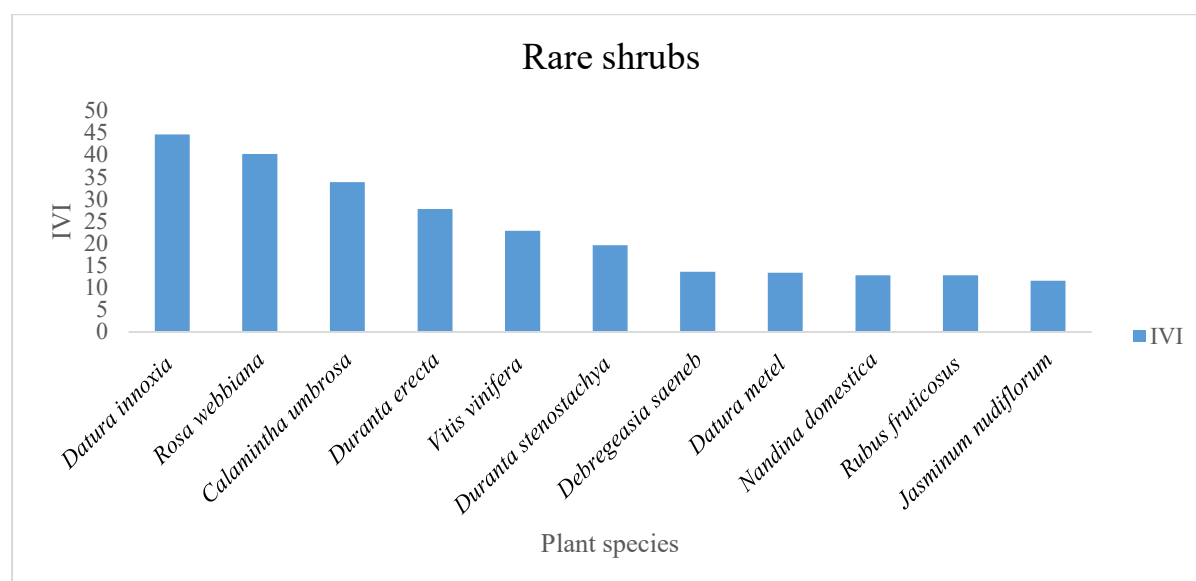


Figure 4.6 The Rare shrub species around the Aik-stream, Sialkot, Pakistan.

#### 4.3.3.3 Abundant and rare herbs

The top most abundant herb species were led by *Parthenium hysterophorus*, with IVI of 1308; following closely were *Cynodon dactylon* (IVI of 762.44), *Eclipta alba* (IVI of 664.40), *Coronops didymus* (IVI of 638.21), *Rumex dentatus* (IVI of 627.42), *Malva neglecta* (IVI of 525.42), *Persicaria glabra* (IVI of 466.16), *Ageratum conyzoides* (IVI of 444.90),



*Dysphania ambrosioides* (IVI of 436.69), and *Sisymbrium irio* (IVI of 387.85) as shown in Figure 3.6. In contrast, less abundant herb species exhibited considerably lower IVI values, among them were *Goldbachia laevigata* (IVI of 5.01), *Paspalum paspalodes* (IVI of 4.865), *Datura innoxia* (IVI of 4.8335), *Sinapis arvensis* (IVI of 4.8309), *Ipomoea carnea* (IVI of 4.80), *Persicaria maculosa* (IVI of 3.976), *Erigeron canadensis* (IVI of 3.90), and *Verbena bonariensis* (IVI of 3.90) as illustrated in Figure 4.7.

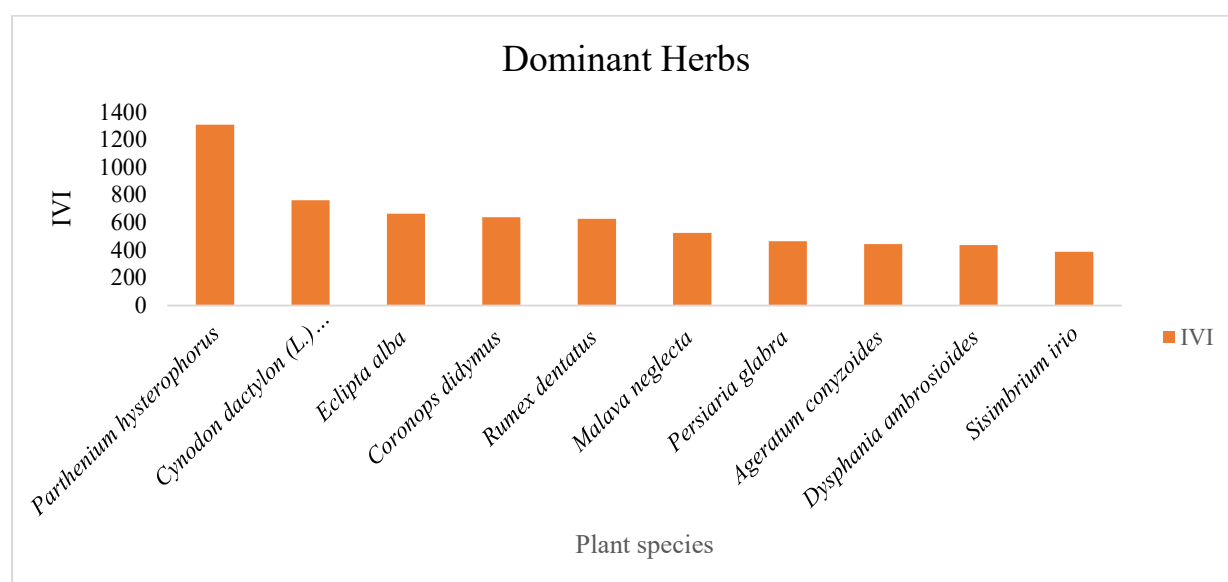


Figure 4.7 The top ten most abundant herb species reported from the Aik-stream area in Sialkot, Pakistan

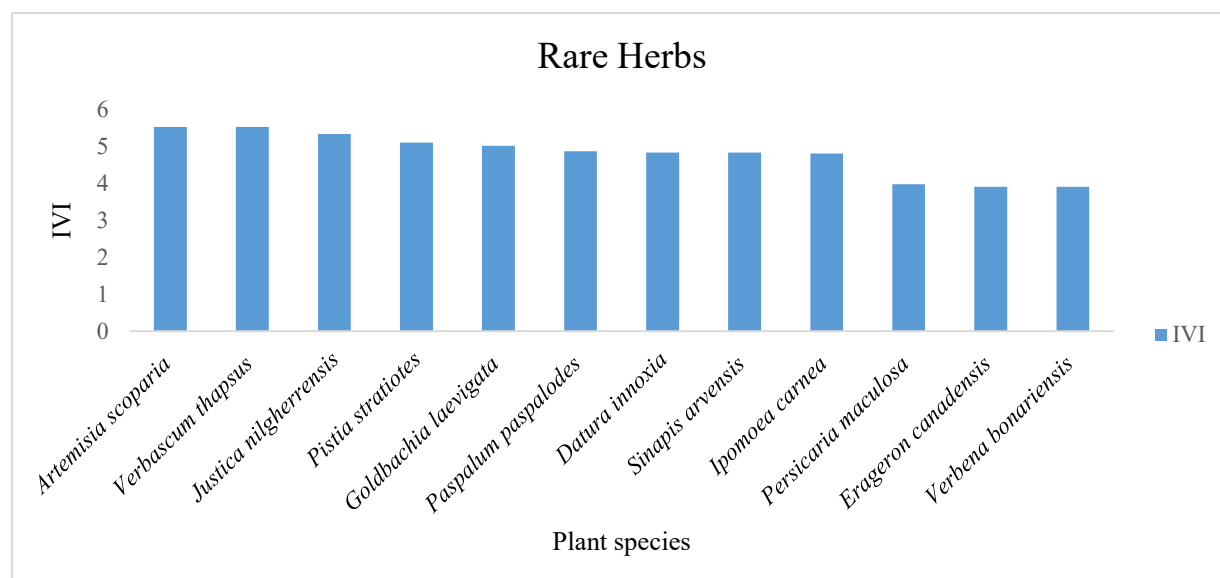


Figure 4.8 Rare herb species with minimum IVI in the studied Aik-stream of Sialkot, Pakistan



#### 4.3.4 Vegetation Classification

The Aik-stream region was divided into three major vegetation zones, namely, the Less Polluted Zone (LPZ), High Polluted Zone (HPZ), and Moderately Polluted Zone (MPZ), using hierarchical Cluster Analysis (CA) and Two-way Cluster Analysis (TWCA) with the Ward Method, which minimizes increases in the sum of squared errors, and Sorenson distance in PC-ORD software. This analysis successfully classified all the stations and plants. Furthermore, it provided insights into the distribution of each plant species at specific stations and even at the quadrat level within different zones, as illustrated in Figures 4.9 and 4.10.

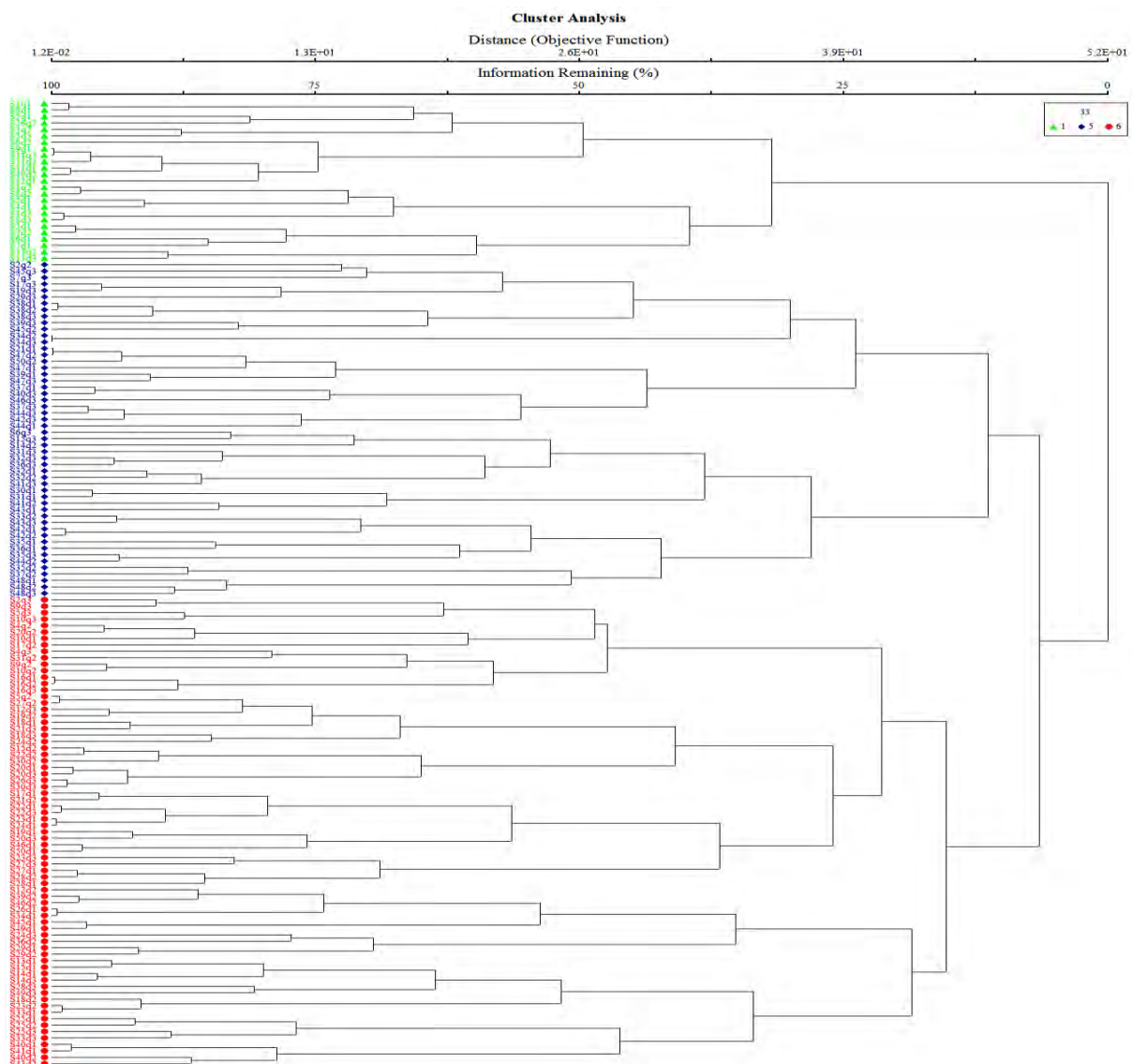


Figure 4.9 Cluster Analysis (CA) of the Aik-stream dividing the region into three distinct zones: the Less Polluted Zone (LPZ), High Polluted Zone (HPZ), and Moderately Polluted Zone (MPZ)



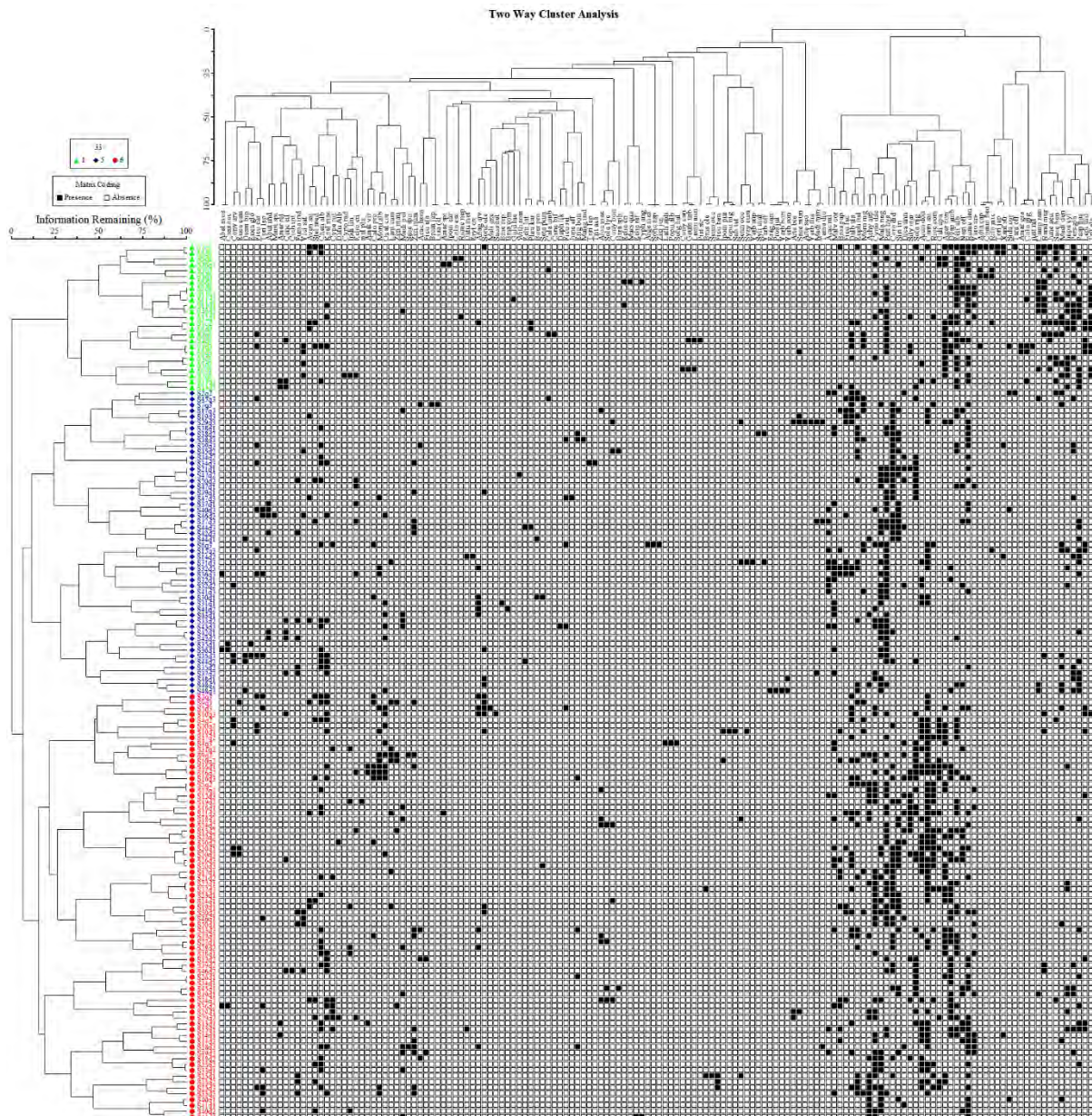


Figure 4.10 The TWCA dendrogram depicted the distribution of plant species near Aik-Stream using Sorensen Distance & Ward Linkage (with narrow single-spaced width).

#### 4.3.5 Species Richness and Diversity Indices

The species richness values range from 238 to 159 plant species in these three major vegetation zones of the Aik-Stream region. The highest Shannon Diversity Index was determined for LPZ ( $H'$  2.035), followed by HPZ and MPZ (1.93). The Simpson Index values, i.e., 0.818, 0.816, and 0.823, were determined for LPZ, HPZ, and MPZ, respectively. At the same time, the Evenness was recorded between 0.726 and 0.814. Other vegetation indices of the Aik-Stream region can be seen in Table 4.3.



Table 4.3 Diversity indices of all the three vegetation zones of the Aik-Stream.

Diversity Indices	LPZ	HPZ	MPZ
Species Richness	238.3091	196.9545	131.3214
Dominance-D	0.181309	0.183901	0.176336
Simpson_1-D	0.818691	0.816098	0.823664
Shannon- H'	2.035018	1.932848	1.93
Evenness-e <sup>H</sup> /S	0.726416	0.814442	0.864775
Brillouin	1.7988	1.687274	1.623214
Menhinick	0.729622	0.660968	0.733368
Margalef	1.858855	1.554262	1.515836
Equitability-J	0.859073	0.902865	0.927204
Fisher-alpha	2.458473	2.030174	2.011757
Berger-Parker	0.311258	0.29475	0.283625
Chao-1	11.09091	9.166667	8.321429
iChao-1	11.09091	9.166667	8.321429
ACE	11.09091	9.166667	8.321429

#### 4.3.5 Indicator of Vegetation zones of the Aik-Stream region

The detailed description of each vegetation zone of the Aik-Stream and its distinctive indicator plants are as follows:

##### 4.3.5.1 Less Polluted Zone (LPZ)

Indicator species analysis revealed the topmost indicators of the Less Polluted Zone (LPZ) in the Aik-stream region. These key indicators, with an indicator value of  $\geq 20\%$  and a probability value of  $\leq 0.05$  after Indicator Species Analysis (ISA), included *Euphorbia helioscopia*, *Dalbergia sisso*, and *Trifolium alexandrinum* (Table 4.3). These species were associated with higher Potassium (K) levels, lower Phosphorus (P) levels, and a neutral soil pH in the LPZ zone (Table 3.3). Additional indicators for this LPZ zone included *Trifolium alexandrinum*, *Koelaria macarantha*, *Silybum marianum*, *Desmostachy bippinanta*, *Morus nigra*, *Nastrum officinales*, *Ranunculus muricatus*, *Paspalum paspalodes*, *Koelaria macarantha*, and *Ranunculus muricatus*, all with  $IV \geq 20\%$  and probability  $\leq 0.05$  (Table 4.4). These indicators were strongly influenced by soil physicochemical conditions, as extensively discussed in Chapter 2.



Table 4.4 The topmost indicator species in relation to significant measured variables along with respective indicator value (IV), probability (p-value) and total importance value index in the Less Polluted Zone

S. No.	Indicator Species	Variable	Max Grp	IV	p-value	TIVI
1	<i>Euphorbia helioscopia</i>	Phosphorus (P)	1	41.7	0.0096	1087.37
2	<i>Dalbergia sisso</i>	pH	6	71.4	0.0006	1116.60
3	<i>Trifolium alexandrinum</i>	Potassium (K)	3	31.2	0.0198	152.06

#### 4.3.5.2 High Polluted Zone (HPZ)

The top three indicator species in the High Polluted Zone (HPZ) were identified as *Cannabis sativa*, *Riccinus communis*, and *Saccharum bengalensis*. These species were indicative of elevated levels of Cadmium (Cd), medium levels of Zinc (Zn), high levels of Chromium (Cr), and low levels of nickel (Ni) when compared to other zones (Table 4.3). Additional characteristic species within this vegetation zone included *Achyranthes aspera*, *Arundo donax*, *Calotropis procera*, *Chenopodium album*, *Coronops didymus*, *Cynodon dactylon*, *Dysphania ambrosioides*, *Eclipta Alba*, *Ficus carica*, *Koelaria macarantha*, *Malva neglecta*, *Parthenium hysterophorous*, *Persicaria glabra*, and *Ziziphus nummularia* (Table 4.5). The indicators in this region are influenced by the soil properties discussed in detail in the second chapter.

Table 4. 5 The topmost indicator species with respect to significant environmental factors, total importance value index, indicator and probability values of High Polluted Zone.

S. No.	Indicator Species	Variable	Max Grp	IV	p-value	TIVI
1	<i>Cannabis sativa</i>	Cr	2	41.3	0.0604	1977.251756
2	<i>Riccinus communis</i>	Ni	2	23.3	0.0912	215.7853242
		Cr	2	46.7	0.0582	
3	<i>Saccharum bengalensis</i>	Cd	3	52.8	0.0324	1087.370032
		Zn	2	33.8	0.03	



#### 4.3.5.3 Moderately Polluted Zone

Following Indicator Species Analysis (ISA), the primary plant indicators of the Moderately Polluted Zone (MPZ) were identified as *Lantana camara*, *Melia azedarach*, and *Silybum marianum* (Table 3.5). These species were associated with higher levels of Copper (Cu), lower levels of Zinc (Zn), and moderate levels of Lead (Pb) within this region (Table 4.6). Additional characteristic species in this zone included *Daphane macronata*, *Euphorbia hirta*, *Geum urbanum* L., *Lantana camara*, *Melia azedarach*, *Rumex hastatus*, *Silybum marianum*, *Sonchus oleraceus*, *Amaranthus viridis*, *Artemisia brevifolia*, *Cenchrus biflorus* Roxb. *Datura innoxia*, and *Erigeron canadensis* (Table 4.7). These indicators respond to changes in environmental factors; a detailed discussion of the soil properties in this region can be found in chapter 2

Table 4.6 The foremost three indicators' species along with significant environmental variables and their respective indicator & probability values and total importance value index in the moderately polluted zone.

S. No.	Indicator Species	Variable	Max Group	IV	p-value	TIVI
1	<i>Lantana camara</i>	Cu	3	45	0.0334	150.583
2	<i>Melia azedarach</i>	Cu	3	30.6	0.0274	231.842
		Zn	1	25	0.0438	
3	<i>Silybum marianum</i>	Pb	2	46	0.0188	198.268

Table 4.7 Other indicator species of the three identified zones of the Aik-Stream region, Sialkot, Pakistan.

Zone 1 (LPZ)					
Indicator species	Variables	Max group	indicator value	p-value	T-IVI
<i>Trifolium alexandrinum</i>	K	3	31.2	0.0198	152.06
	P	2	39.3	0.0236	
<i>Koelaria macarantha</i>	P	2	45.2	0.021	135.64



<i>Silybum marianum</i>	Zn	1	44	0.0078	153.81
<i>Dalbergia sisso</i>	pH	6	71.4	0.0006	1116.60
<i>Desmostachy bippinanta</i>	pH	6	25	0.0806	
<i>Morus nigra</i>	pH	6	43.5	0.0168	352.08
<i>Nastrum officinales</i>	pH	6	36.3	0.0446	200.09
<i>Ranunculus muricatus</i>	P	2	34.4	0.0292	125.33
<i>Paspalum paspalodes</i>	P	2	31.4	0.044	177.89
<i>Trifolium alexandrinum</i>	P	2	31.2	0.0198	102.06
	K	3	45.2	0.021	
<i>Koelaria macarantha</i>	P	2	58.7	0.0074	135.64
<i>Ranunculus muricatus</i>	P	1	29	0.071	125.33
<i>Euphorbia helioscopia</i>	P	1	41.7	0.0096	1087.37
<b>Zone 2 (HPZ)</b>					
	<b>Variable</b>	<b>Max group</b>	<b>indicator value</b>	<b>p-value</b>	<b>T-IVI</b>
<i>Achyranthes aspera</i>	Cr	6	47.1	0.0542	242.41
	Cd	3	50	0.0408	
<i>Arundo donax</i>	Zn	4	38	0.018	939.05
	Cd	3	33.2	0.0788	
<i>Calotropis procera</i>	Cr	9	40.1	0.0562	228.85
<i>Cannabis sativa</i>	Cr	2	41.3	0.0604	1977.25
<i>Chenopodium album</i>	Cu	2	42.7	0.036	273.55
<i>Coronops didymus</i>	Cd	3	50	0.0374	374.52



<i>Cynodon dactylon</i>	Pb	4	42.1	0.0554	336.08
<i>Dysphania ambrosioides</i>	Cu	4	81.4	0.0008	298.50
<i>Eclipta Alba</i>	Cr	2	43.5	0.0688	398.99
<i>Ficus carica</i>	Zn	2	29	0.071	420.11
<i>Koelaria macarantha</i>	Pb	4	50	0.0406	137.94
<i>Malava neglecta</i>	Zn	6	37.4	0.0184	299.04
<i>Riccinus communis</i>	Cr	2	46.7	0.0582	215.78
	Ni	2	23.3	0.0912	
<i>Parthenium hysterophorous</i>	Pb	1	33.6	0.0492	636.72
	Cd	3	46.7	0.0434	
<i>Persiaria glabra</i>	Cd	3	41.7	0.0096	203.98
	Cr	2	46.9	0.0186	
<i>Saccharum bengalensis</i>	Cd	3	52.8	0.0324	1087.37
	Zn	2	33.8	0.03	
<i>Ziziphus nummularia</i>	Cu	2	51.7	0.0456	313.49
<b>Zone 3 (MPZ)</b>					
	<b>Variable</b>	<b>Max group</b>	<b>indicator value</b>	<b>p *</b>	<b>T-IVI</b>
<i>Daphane macronata</i>	Zn	1	52	0.0308	192.52
<i>Euphorbia hirta</i>	Pb	2	53.7	0.057	116.27
	Zn	1	47.9	0.0688	
<i>Geum urbanum L.</i>	Cu	3	49.2	0.0224	191.35
	Pb	2	52.6	0.0274	
<i>Lantana camara</i>	Cu	3	45	0.0334	150.58
<i>Melia azedarach</i>	Cu	3	30.6	0.0274	231.84
	Zn	1	25	0.0438	
<i>Rumex hastatus</i>	Cu	3	37	0.0356	112.79
	Pb	2	39	0.041	



<i>Silybum marianum</i>	Pb	2	46	0.0188	198.26
<i>Sonchus oleraceus</i>	Zn	1	52	0.0422	115.83
<i>Amaranthus viridis</i>	Cd	2	44.1	0.041	140.39
	Pb	2	52	0.0306	
<i>Artemisia brevifolia</i>	Cu	3	52	0.0454	131.88
	Pb	2	35.4	0.0306	
<i>Cenchrus biflorus</i> Roxb.	Pb	2	35.2	0.0094	137.97
	Zn	1	54.8	0.0306	
<i>Datura innoxia</i>	Cu	3	52	0.0562	122.35
	Pb	2	48.7	0.0346	
	Zn	1	52	0.0144	
<i>Erigeron canadensis</i>	Cu	3	36	0.0421	99.41

#### 4.3.6 Direct Ecological Gradient through Canonical Correspondence Analysis (CCA)

CCA again confirmed the plant indicators of different vegetation zones of Aik Stream. The direct ecological gradient analysis showed that the precipitation, temperature and soil (pH, EC, TDS, P, K, Ni, Cd, Cr, Cu, Pb, As, Hg, Zn) have a significant impact ( $p = 0.0001$ ) on indicator plants diversity of the Aik stream region (Table 4.8). For example, indicator species of the LPZ were clustered under the influence of higher potassium, medium phosphorus, and alkaline pH. Meanwhile, indicator species of HPZ were under the impact of a moderate amount of precipitation, temperature, nearly neutral soil pH, and higher amounts of heavy metals such as Cr, Cd, Ni, and As compared to other zones. The indicators of MPZ were low precipitation, higher temperature, Pb, Zn, and Cu compared to other zones.



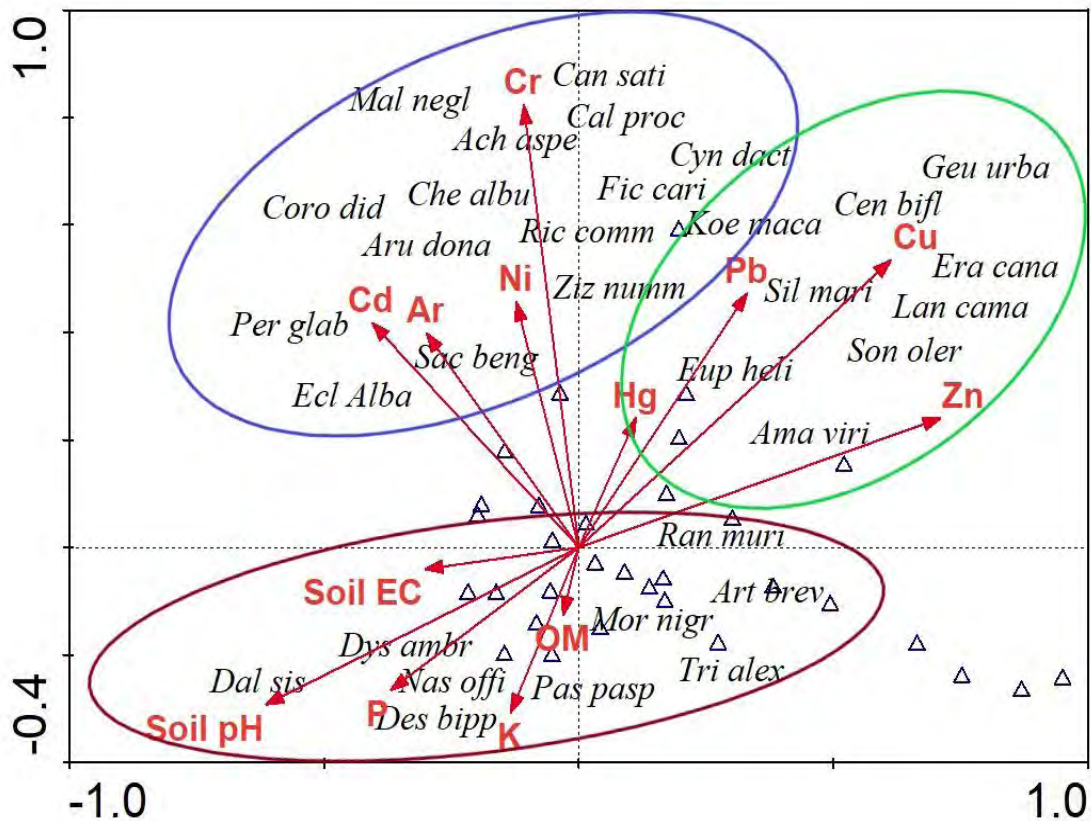


Figure 4.11 Indicator species of all the studied zones with their influencing environmental factors through Canonical Correspondence Analysis (CCA).

#### 4.3.7 Analyzing the Impact of Environmental Factors on Plant Abundance in different polluted zones of Aik Stream, Sialkot, Pakistan Using GLM

##### 4.3.7.1 Less Polluted Zone (LPZ)

Generalized Linear Model (GLM) was employed to analyze the influence of various environmental and meteorological factors on the Importance Value Index (IVI) of plants in the Less Polluted Zone (LPZ) of Aik Stream in Sialkot, Pakistan. The variables examined included temperature, precipitation, and several environmental parameters such as pH, organic matter, phosphorus (P), potassium (K), and various heavy metals (Cu, Zn, Cr, Pb, Cd, Ni, As, Hg). Findings showed that temperature and precipitation positively and significantly affect plant IVI, with coefficients of 0.729 and 0.182, respectively. In contrast, the impact of pH, P, K, Cu, and As was negative and significant, with respective coefficients of 0.713, 0.124, 0.209, 0.556, and 0.506. Among these factors, temperature had the most substantial effect, with a coefficient of 0.97, suggesting that a one-degree increase in temperature could lead to a 97-fold increase in expected IVI. Similarly, a unit change in precipitation was



associated with an 89-fold increase in expected IVI. The model's fit was considered good, as indicated by various statistical measures ( $R^2 = 0.442$ , AIC = 1123, BIC = 1043, DF = 121, Chi-Square = 11.21, and p-value = 0.033). The study found no evidence of multicollinearity among the explanatory variables, as all variance inflation factor (VIF) values were below 5. These results contribute to a deeper understanding of how environmental conditions affect plant communities.

Table 4.8 Generalized linear model assessing the influence of environmental factors on plant abundance in the LPZ locality of Aik Stream vegetation. IVI serves as the dependent variable.

Independent variables	Coefficient ( $\beta$ )	SE	Exp ( $\beta$ )	t-value	p-value
(Intercept)	<b>0.969**</b>	21.43884	0.133	2.24	0.025
Temp	<b>0.729**</b>	5.749441	0.974	2.39	0.017
Precipitation	<b>0.182*</b>	0.239998	0.894	2.76	0.047
pH	<b>-0.713***</b>	1.858153	1.013	-3.08	0.009
EC	-0.026	0.148267	0.976	-0.18	0.859
TDS	-0.111	0.096342	3.352	-1.16	0.247
OM	0.0136	0.083203	0.573	0.16	0.878
P	<b>-0.124***</b>	0.790336	0.735	-3.03	0.006
K	<b>-0.209***</b>	0.770375	0.981	-2.57	0.004
Cu	<b>-0.556*</b>	0.298592	1.1757	-2.86	0.062
Zn	-0.307	0.273128	0.969	-1.13	0.264
Cr	-0.018	0.377585	1.044	-0.05	0.961
Pb	-0.161	0.164038	1.660	0.99	0.324
Cd	-0.031	0.061888	0.833	-0.5	0.616
Ni	-0.043	0.105906	0.133	0.41	0.679
As	<b>-0.506**</b>	0.215159	0.974	2.36	0.018
Hg	-0.181	0.149551	0.894	-1.22	0.224

Table 4.9 Indicate the model fit value for GLM.

Model fit	$R^2$	AIC	BIC	DF	Chi Squ	P value
GLM	0.442	1123.34	1043.23	33	11.21	0.033

#### 4.3.7.2 High Polluted Zone (HPZ)

The influence of environmental factors on the Importance Value Index (IVI) of plant species in the HPZ area was assessed. The analysis revealed that temperature positively and significantly affects plant IVI (0.844), while precipitation exhibits a negative and significant impact (-0.019). Additionally, the presence of heavy metals (Cu, Zn, Cr, Pb, Cd, Ni, As, Hg) in the Aik Stream of the HPZ area showed a negative and significant correlation with plant IVI, with coefficients of 0.126, 0.120, 0.330, 0.239, 0.120, 0.341, 0.153, and 0.049,



respectively. The exponential beta ( $\exp(\beta)$ ) coefficients indicate a multiplicative effect of each predictor on IVI. Temperature emerged as a strong predictor with a significant estimate of 0.885, suggesting that a one-degree increase in temperature correlates with an 88-fold increase in the expected IVI. Contrary to initial findings, the revised analysis shows that precipitation positively and significantly affects IVI (0.981), meaning a unit increase in precipitation leads to a 98-fold increase in expected IVI. The model's fitness was evaluated using various statistical measures ( $R^2 = 0.621$ , AIC = 1240, BIC = 1156, DF = 133, Chi-Square = 14.23, p-value = 0.022), confirming its suitability for a General Linear Model (GLM). The study also confirmed the absence of multicollinearity among explanatory variables, as evidenced by the variance inflation factor test, where all values were below 5. These findings contribute to a deeper understanding of the dynamic relationship between environmental conditions and plant communities.

Table 4.10 Generalized linear model assessing the influence of environmental factors on plant abundance in the HPZ locality of Aik Stream vegetation. IVI serves as the dependent variable

Independent variables	Coefficient ( $\beta$ )	SE	Exp ( $\beta$ )	t-value	p-value
(Intercept)	0.948	39.13235	0.133	-0.59	0.558
Temp	<b>0.844**</b>	10.16343	0.885	2.97	0.033
Precipitation	<b>-0.019**</b>	0.250821	0.981	-2.08	0.039
Soil pH	-0.925	1.186399	0.145	-3.62	0.005
EC	-0.072	0.082456	0.930	-0.87	0.382
TDS	0.081	0.08539	1.084	0.95	0.346
OM	0.104	0.077936	1.110	1.34	0.185
P	-0.109	0.252241	0.896	-0.43	0.664
K	-0.510	0.319539	0.600	-1.6	0.116
Cu	<b>-0.126**</b>	0.22546	0.973	-3.12	0.006
Zn	<b>-0.120**</b>	0.24855	0.886	-2.49	0.027
Cr	<b>-0.330**</b>	0.151179	0.718	-2.19	0.029
Pb	<b>-0.239***</b>	0.236829	1.009	-3.04	0.007
Cd	<b>-0.120**</b>	0.125963	1.128	-2.96	0.038
Ni	<b>-0.341**</b>	0.082127	1.042	-2.597	0.014
As	<b>-0.153**</b>	0.122666	0.857	-2.25	0.011
Hg	<b>-0.049**</b>	0.07847	1.050	-2.63	0.029

Table 4.11 indicates the model fit value for GLM.

Model fit		$R^2$	AIC	BIC	DF	Chi Squ	P value
GLM		0.621	1240.34	1156.23	49	14.23	0.022



#### 4.3.7.3 Moderate Polluted Zone (MPZ)

The impact of environmental variables on plant abundance in MPZ found that temperature and precipitation positively and significantly influence plant IVI (0.883), while electrical conductivity (EC), copper (Cu), zinc (Zn), and lead (Pb) negatively and significantly impact plant IVI in the Aik Stream MPZ area (0.40, 0.219, 0.717, 0.155 respectively). The exponential coefficients ( $\exp(\beta)$ ) suggest a multiplicative effect of each predictor on IVI. Temperature was identified as having a strong, significant influence, with an estimate of 0.231. This indicates that a one-degree increase in temperature corresponds to a 32-fold increase in expected IVI. Similarly, precipitation has a positive, significant impact on IVI (0.7908), implying that a one-unit increase in precipitation results in a 79-fold increase in expected IVI. The model's fit, presented in Table 5, is considered good for Generalized Linear Models (GLM), with the following statistics:  $R^2 = 0.355$ , AIC = 959, BIC = 865, DF = 133, Chi-Square = 10.23, and a p-value of 0.0432. The study found no multicollinearity issues among the explanatory variables, as indicated by variance inflation factor tests, all values of which were below 5. These findings contribute to a deeper understanding of the intricate relationship between environmental conditions and plant communities.

Table 4.12 Generalized linear model assessing the influence of environmental factors on plant abundance in the MPZ locality of Aik Stream vegetation. IVI serves as the dependent variable.

Independent variables	Coefficient ( $\beta$ )	SE	Exp ( $\beta$ )	t-value	p-value
(Intercept)	0.675	39.60012	0.321	-1.23	0.219
Temp	<b>0.883***</b>	9.76181	0.7908	3.63	0.004
Precipitation	<b>0.769**</b>	0.449507	2.1576	2.71	0.047
Soil pH	-1.808	2.048602	0.1638	-0.88	0.377
EC	<b>-0.400*</b>	0.217635	0.6702	-1.84	0.066
TDS	0.197	0.128648	1.2185	1.54	0.124
OM	0.031	0.138085	1.0322	0.23	0.818
P	-0.380	0.275152	0.6832	-1.38	0.166
K	-0.371	0.575945	0.6899	-0.64	0.519
Cu	<b>-0.219*</b>	0.34107	1.2453	-2.64	0.052
Zn	<b>-0.717**</b>	0.336441	0.4877	-2.13	0.033
Cr	-0.267	0.185177	0.7655	-1.44	0.149
Pb	<b>-0.155**</b>	0.263686	0.8555	-2.59	0.050
Cd	0.139	0.058398	1.1493	-2.38	0.017
Ni	0.081	0.183863	1.0846	0.44	0.659
As	0.293	0.238815	1.3410	1.23	0.219
Hg	0.013	0.178456	1.0132	0.07	0.941



Table 4. 13 Indicate the model fit value for GLM.

Model fit	R <sup>2</sup>	AIC	BIC	DF	Chi Squ	P value
GLM	0.355	959.34	865.23	16	10.23	0.0432

#### 4.3.8 Correlation between Alien and Invasive Species with Native Species Diversity

The relations between alien and invasive species richness and abundance and native species diversity varied with the intensity of invasion. For alien species, in plots with low alien cover (RAC < 5%), there were positive correlations between alien species richness ( $S_{\text{alien}}$ ) and native species richness ( $S_{\text{native}}$ ), as well as between RAC and both native evenness ( $J_{\text{native}}$ ) and native diversity ( $H'_{\text{native}}$ ). However, in plots with higher alien cover (RAC > 5%), negative correlations emerged between  $S_{\text{alien}}$  and  $J_{\text{native}}$ , and between RAC and both  $S_{\text{native}}$  and  $H'_{\text{native}}$  (Table 4.14).

Similarly, for invasive species, in plots with low invasive cover (RIC < 5%), there were negative correlations between invasive species richness ( $S_{\text{invasive}}$ ) and native species richness ( $S_{\text{native}}$ ), but positive correlations between  $S_{\text{invasive}}$  and native diversity ( $H'_{\text{native}}$ ), and between RIC and both  $J_{\text{native}}$  and  $H'_{\text{native}}$ . In plots with higher invasive cover (RIC > 5%), negative correlations were noted between  $S_{\text{invasive}}$  and  $J_{\text{native}}$ , and between RIC and both  $S_{\text{native}}$  and  $H'_{\text{native}}$ . These findings indicate that the existence and profusion of alien and invasive species can significantly impact native species diversity, with the nature of these impacts varying depending on the level of cover by alien and invasive species

**Table 4.14** Spearman's correlation between native diversity and alien/invasive species richness with relative cover, p value in parentheses.

Native diversity	RAC < 5% (n = 26)		RAC > 5% (n = 124)	
	S alien	RAC	S alien	RAC
$S_{\text{native}}$	0.245 (0.252)	0.174 (0.027)	-0.132 (0.003)	-0.464 (0.0049)
$J_{\text{native}}$	0.235 (0.574)	0.443 (0.073)	-0.434 (0.034)	-0.376 (0.026)
$H'_{\text{native}}$	0.352 (0.021)	0.421 (0.047)	-0.243 (0.052)	-0.434 (0.014)
	RIC < 5% (n = 84)		RIC > 5% (n = 65)	
	S Invasive	RIC	S Invasive	RIC
$S_{\text{native}}$	-0.163 (0.052)	-0.372 (0.062)	-0.563 (0.034)	-0.632 (0.029)
$J_{\text{native}}$	0.734 (0.317)	0.382 (0.049)	-0.295 (0.087)	-0.442 (0.048)
$H'_{\text{native}}$	0.253 (0.047)	0.726 (0.047)	-0.472 (0.145)	-0.382 (0.367)

S = species richness (species/plot), J Pielou evenness, H' Shannon diversity



#### 4.3.9 Factors Shaping Species Composition

The Mantel tests (Table 4.15) identified notable correlations ( $r = 0.3592$ ,  $P \leq 0.01231$ ) between matrices that outline environmental attributes, disturbance, metal contamination, soil characteristics, and overall composition of native, alien, and invasive species).

Significant relationships were found between plant matrices and the Metal Pollution Index (MPI), with higher correlations for invasive species ( $r = 0.834$ ) and alien species ( $r = 0.745$ ), and slightly weaker relations for native species ( $r = 0.453$ ), associations ( $r = 0.387$ ), and alliances ( $r = 0.264$ ). The relationships between the disturbance matrix and native species were quite weak, but strong correlations were observed for alien species ( $r = 0.654$ ) and invasive species ( $r = 0.554$ ). Soil properties were strongly correlated with native species composition ( $r = 0.563$ ) and overall species composition ( $r = 0.674$ ). Notably, Plots with similar environmental settings did not reflect similar disturbance conditions.

**Table 4.15** Standardized Mantel correlation statistic ( $r$ ) between dissimilarity matrices (Bray-Curtis distance) for all species ( $n = 216$ ), native species ( $n = 120$ ), non-native species ( $n = 96$ ), and invasive species ( $n = 32$ ), environmental variables, soil properties, Metal Polluted Index (MPI) and Disturbance Index (DI).

	Native species	Alien species	Invasive species	All species	Environment	MPI	Soil
<b>Alien species</b>	0.287 (0.0065)						
<b>Invasive species</b>	0.432 (0.0036)	0.765 (0.041)					
<b>Environment</b>	0.198 (0.008)	0.036 (0.007)	0.56 (0.0027)	0.146 (.0047)			
<b>DI</b>	0.423 (0.0015)	0.654 (0.0047)	0.554 (0.0029)	0.231 (0.0019)	0.237 (0.0022)		
<b>MPI</b>	0.453 (0.0023)	0.745 (0.0048)	0.834 (0.0072)	-0.634 (0.0352)	0.342 (0.005)	0.43 (0.03)	
<b>Soil</b>	0.563 (0.0043)	0.321 (0.009)	0.421 (0.0027)	0.674 (0.001)	0.365 (0.0043)	0.32 (0.01)	0.573 (0.0035)

Comparisons derived from sample plot values ( $n = 150$ ); for comparisons involving the alien species matrix,  $n = 96$ . Refer to Table 1 for the environmental and disturbance variables and Fig. 2 for details on alliances and associations.



#### 4.4 Discussion

Our study specifically focuses on the flora of ecosystems impacted by industrial wastewater pollution, examining the effects of environmental factors on species composition and distribution patterns around the Aik-Stream region in Sialkot, Pakistan. Our research identified 182 plant species belonging to 136 genera and 49 families in the areas around Aik-Stream, which vary in pollution levels. These species are categorized into 127 herbs, 26 shrubs, and 28 trees. Notably, the Poaceae family predominates, followed by Asteraceae, Fabaceae, and Moraceae. This finding aligns with studies by (Saarela et al., 2018) and (Majeed et al., 2022), who also reported the dominance of the Poaceae family in similar regions. The prevalence of Poaceae is attributed to its wide ecological adaptability, human disturbances, and perennating capabilities (Manan et al., 2022). Key plant species observed in our study include *Dalbergia sisso*, *Ficus carica*, *Morus alba*, and several others. These findings are consistent with (A. Nazir, Malik, Ajaib, Khan, & Siddiqui, 2011), who identified similar species in Rawalpindi and Islamabad, Pakistan industrial areas. Moreover, (Noreen et al., 2019) reported the presence of species like *Calotropis procera* and *Ricinus communis* in regions affected by oil and gas pollution. The research by (Lavergne, Mouquet, Thuiller, & Ronce, 2010) emphasizes the significance of interactions between different species in shaping the distribution and diversity of species in various environments. We employed Species Area Curves and Cluster Analyses to analyze the adequacy of our sampled size and the distribution of each plant species. These methods have been widely used in ecological research to assess plant species distribution and sampling adequacy, as demonstrated in studies by (Jiménez-Valverde, Lobo, & Hortal, 2009; Syfert, Smith, & Coomes, 2013).

This chapter is focused on exploring how industrial wastewater pollution and climatic and soil (edaphic) conditions influence the development of various plant communities. Employing multivariate statistical methods, the study examines how these factors contribute to forming distinct vegetation zones. The research utilized ecological software for cluster analysis, which helped categorize vegetation into three different pollution zones: less polluted, highly polluted, and moderately polluted. This classification was based on the similarities in plant species composition and the variations observed in climate, soil conditions, and topography. The study underscores the critical role that these environmental gradients play in shaping vegetation patterns, echoing the findings of previous research by (Rowe, 2009), who also highlighted the significance of these factors in determining plant community composition.



The Less Polluted Zone (LPZ), with its specific species thriving under conditions of higher precipitation and distinct soil nutrient levels, exemplifies the delicate balance of nature and how subtle changes in environmental conditions can significantly influence plant communities. Applying statistical methods like GLM, CCA, and SEM reiterates this point, offering a nuanced understanding of the interdependencies between climatic, edaphic, and pollution factors. The High Pollution Zone (HPZ), with its unique set of flora, further illustrates this complexity. The presence of indicator species in an environment marked by industrial pollution and specific soil nutrient profiles highlights the resilience and adaptability of certain plants. This zone, showing lower biodiversity, serves as a stark reminder of the potential impacts of environmental stressors on ecological diversity and health.

Similarly, the Moderate Pollution Zone (MPZ) provides another perspective on the interaction between vegetation and environmental pollutants, particularly heavy metals. The comparison with historical studies by (Bayouli et al., 2021) underscores the long-standing recognition of certain plants as bio-indicators of specific soil conditions, particularly mineral deposits. This continuity in ecological research highlights the enduring relevance of understanding plant responses to environmental gradients.

The comprehensive observations made in pollution zones underscore the intricate interplay of various environmental factors i.e., climate, soil nutrients, and industrial pollutants, in shaping the ecological landscape. This detailed understanding is pivotal for formulating effective environmental management strategies to preserve these diverse ecosystems. For instance, *Silene cobalticola*, confined to copper and cobalt mines in Savannah and classified as critically endangered by the IUCN, illustrates how specific plants are uniquely adapted to certain mineral-rich environments. Similarly, plants like *Acalypha dikuluwensis*, *Ocimum centraliafricanum*, and others serve as bioindicators for copper, a practice dating back to the 17th century when Scandinavian miners used *S. suecica* for locating copper deposits. Further emphasizing the role of plants as mineral indicators, *Polycarpha spirostylis* is renowned as a copper indicator, with its reputation established by early researchers (Cole & Smith, 1984) and subsequently validated through field studies.

Additionally, *Agathis ovata*'s growth is influenced by nickel, chromium, and magnesium concentrations, highlighting the diverse responses of vegetation to different metal deposits. Iron indicators such as *Acacia athens* and others, along with *Eriachne mucronata*, identified as an indicator of copper, lead, and zinc mineralization, showcase the varied



ecological adaptations of plants to metal-rich environments. The study also sheds light on other bioindicators like *Hybanthus austrocaledonicus* for nickel and species indicating the presence of selenium and uranium. Brooks' research in 1979 on zinc indicators adds to this extensive list of plant-based environmental indicators. Moreover, contemporary studies like those of (Madkour & Laurence, 2002) in Egypt and (Johnston et al., 2007) have expanded the scope of bioindicators to include plants that respond to atmospheric conditions like ozone and a wide range of hydrogeomorphic and soil conditions. The research also touches upon the remediation capabilities of plants in mining regions, with studies by (Chandra, Saxena, & Kumar, 2015) highlighting the potential of phytoremediation in environmental management. This extensive body of research, like the current study, underscores the critical role of plants not only as indicators of environmental conditions but also as active participants in ecological restoration and management.

The GLM analysis helped elucidate the reasons behind the varying abundance of certain plant species across different pollution gradient zones. Key findings pointed to the significant influence of heavy metals such as chromium, cadmium, nickel, zinc, lead, and copper, along with soil pH, phosphorus, and potassium levels, on plant species abundance and rarity. This is in line with observations made by (Alfeus, Molnar, Boman, Hopke, & Wichmann, 2023), who noted that industrial particulate matter often harbors high concentrations of heavy metals, impacting surrounding water bodies, soil, and vegetation. Plants near pollution sources can absorb these contaminants during nutrient uptake, leading to profound physiological changes, especially in species with limited genetic diversity and narrow ecological niches. The distribution patterns and composition of plant species in the Aik-Stream region, assessed using CCA revealed that the mentioned environmental variables play a crucial role in the distribution of regional plant species. (Kabir, Iqbal, & Shafiq, 2012), conducted related research in a cement-polluted ecosystem in Karachi, Pakistan, finding higher concentrations of calcium carbonate and other soil properties, emphasizing the importance of such studies for better environmental management.

Although the extent of invasion differed in region, exhibited both minimal and substantial coverage of alien species. Plots with minimal alien species coverage ( $RAC < 5\%$ ) were found in all plant associations except for Alliance 1 (midstream/HPZ), while plots with higher alien cover ( $RAC > 5\%$ ) were recorded in all four alliances. For riparian vegetation along the Aik-Stream, the correlations between alien species richness and RAC with native species diversity revealed regions susceptible to invasion, particularly those with high native



diversity and regions where alien species abundance or coverage had a adverse impact on native diversity (e.g.,  $RAC > 5\%$ ). In streamside ecosystems, the richness of alien species is clearly associated with native species richness (Brummer et al. 2016, Peng et al. 2019, Tomasetto et al. 2019).

In plots with low levels of alien coverage ( $RAC < 5\%$ ) along the Aik-stream, there was a direct association between the richness of alien species ( $S_{alien}$ ) and both the richness of native species ( $S_{native}$ ) and the diversity of native species ( $H'_{native}$ ). Moreover,  $RAC$  rose in conjunction with higher native evenness ( $J_{native}$ ) and diversity ( $H'_{native}$ ), lending support to the 'rich get richer' hypothesis, which suggests that more diverse communities may create conditions that favour the establishment of alien species (Carpio et al. 2017, McKinney and Kark 2017, Duncan et al. 2019). These direct correlations could indicate early stages of invasion or coexistence with less competitive invaders (Cubino et al. 2022).

In contrast, in areas with high alien cover ( $RAC > 5\%$ ), no correlation was observed between alien species richness ( $S_{alien}$ ) and native species richness ( $S_{native}$ ). As  $RAC$  increased, both native species richness ( $S_{native}$ ) and diversity ( $H'_{native}$ ) decreased, and a weak negative correlation was observed between alien species richness ( $S_{alien}$ ) and native evenness ( $J_{native}$ ). Similar trends were observed for invasive species. This suggests that higher alien cover is associated with negative impacts on native species diversity, likely due to competitive replacement of native species by alien and invasive species (Carboni et al. 2021, Rojas-Sandoval et al. 2022, Vujanović et al. 2022).

Our research findings have opened new avenues for understanding the complex interactions between vegetation and environmental pollution. We have gained valuable insights into how different pollutants influence vegetation in varied ecological settings by employing a comprehensive suite of multivariate statistical techniques. The findings of this study serve as a critical foundation for further research, which should aim to expand its scope geographically and methodologically. Future research, guided by the recommendations proposed, should focus on long-term monitoring to capture the temporal dynamics of pollution impact, explore the potential of phytoremediation, and delve deeper into the genetic adaptations of plants in polluted environments. Integrating advanced technological tools and considering more comprehensive environmental variables will enhance the precision and depth of subsequent studies. Moreover, understanding the broader ecological implications, including impacts on wildlife and the socio-economic ramifications for dependent



communities, is essential. Collaborative efforts involving policymakers, local communities, and researchers will ensure that the findings are scientifically relevant and practically applicable to environmental management and conservation strategies. As this research outlines, the path forward is deepening our ecological understanding and actively contributing to sustainable environmental practices and policies. Our study underscores the necessity of a multidisciplinary approach in addressing the challenges posed by environmental pollution, aiming ultimately to preserve and restore the delicate balance of our ecosystems.

#### 4.5 Conclusion

The research conducted in the Aik-Stream area of Sialkot, Pakistan, has led to significant findings regarding the region's plant biodiversity and the impact of environmental factors on vegetation patterns. This area is characterized by a rich diversity of plant species, with 182 distinct species identified, among which the Family Poaceae emerges as the most dominant in the pollution-affected areas. Key species such as *Ficus carica*, *Calotropis procera*, *Cynodon dactylon*, *Dalbergia sisso*, *Saccharum bengalensis* etc. are noted for their prevalence, while *Populus nigra*, *Verbena bonariensis*, and *Nerium oleander* are categorized as rarer species. Through General Linear Model (GLM) analysis, the study underscores the substantial influence of various environmental variables on the occurrence of both dominant and rare plant species within this region. Concurrently, Canonical Correspondence Analysis (CCA) reveals that factors such as soil pH, temperature, precipitation, phosphorous, potassium, and heavy metals play a crucial role in determining the distribution patterns of plant species. The study also highlights the substantial effects of water pollution, climatic conditions, and soil characteristics on the overall vegetation structure of the Aik-stream area. Indicator Species Analysis (ISA) has proven to be an exceptionally effective technique for identifying indicator plants, a process further validated through CCA, GLM, and Structural Equation Modeling (SEM) analyses. This approach, being applied for the first time, allows these indicators to be utilized for various purposes, including reforestation initiatives and innovative habitat plantation strategies. SEM analysis, in particular, has shed light on the complex interplay between measured environmental factors and vegetation dynamics in the Aik-Stream area. The methodology employed in this study, encompassing statistical and modeling techniques, offers a replicable framework for classifying vegetation and identifying indicator plants in any geographical region or microhabitat worldwide. The next chapters will



focus on monitoring indicator plants to evaluate the potential of phytoremediation and delve deeper into the adaptations of plants in polluted environments.





#### OPEN ACCESS

EDITED BY  
Mohiuddin Md. Taimur Khan,  
Washington State University Tri-Cities,  
United States

REVIEWED BY  
Keith Dana Thomsen,  
Lawrence Livermore National  
Laboratory (DOE), United States  
Li Chen,  
Lanzhou University, China

\*CORRESPONDENCE  
António Raposo  
antonio.raposo@ulusofona.pt  
Heesup Han  
heesup.han@gmail.com

SPECIALTY SECTION  
This article was submitted to  
Planetary Health,  
a section of the journal  
Frontiers in Public Health

RECEIVED 02 August 2022  
ACCEPTED 26 September 2022  
PUBLISHED 13 October 2022

CITATION  
Ejaz U, Khan SM, Aqeel M, Khalid N,  
Sarraz W, Naeem N, Han H, Yu J,  
Yue G and Raposo A (2022) Use of  
*Parthenium hysterophorus* with  
synthetic chelator for enhanced  
uptake of cadmium and lead from  
contaminated soils—a step toward  
better public health.  
Front. Public Health 10:1009479.  
doi: 10.3389/fpubh.2022.1009479

# Use of *Parthenium hysterophorus* with synthetic chelator for enhanced uptake of cadmium and lead from contaminated soils—a step toward better public health

Ujala Ejaz<sup>1</sup>, Shujaul Mulk Khan<sup>1,2</sup>, Muhammad Aqeel<sup>3</sup>,  
Noreen Khalid<sup>4</sup>, Wajiha Sarraz<sup>5</sup>, Nayab Naeem<sup>6</sup>,  
Heesup Han<sup>5\*</sup>, Jongsik Yu<sup>6</sup>, Gong Yue<sup>7</sup> and António Raposo<sup>8\*</sup>

<sup>1</sup>Department of Plant Sciences, Quaid-i-Azam University, Islamabad, Pakistan, <sup>2</sup>Member, Pakistan Academy of Sciences, Islamabad, Pakistan, <sup>3</sup>State Key Laboratory of Grassland Agro-ecosystems, College of Ecology, Lanzhou University, Lanzhou, China, <sup>4</sup>Department of Botany, GC Women University, Sialkot, Pakistan, <sup>5</sup>College of Hospitality and Tourism Management, Sejong University, Seoul, South Korea, <sup>6</sup>College of Business Division of Tourism and Hotel Management, Cheongju University, Cheongju-si, South Korea, <sup>7</sup>Business School Tourism and Hospitality Management, Xuzhou University of Technology, Xuzhou City, China, <sup>8</sup>CBIO5 (Research Center for Biosciences and Health Technologies), Universidade Lusófona de Humanidades e Tecnologias, Lisboa, Portugal

*Parthenium hysterophorus* L. is a vigorous plant species with cosmopolitan distribution. It can uptake considerable quantities of heavy metals from the soil and accrue these metals in its different tissue. The use of chelating agent i.e., Ethylenediaminetetraacetic acid (EDTA) can boost up metal uptake capacity. Pot experiment was performed to evaluate phytoextraction potential of *P. hysterophorus* for lead (Pb) and cadmium (Cd) with and without the aid of



## DYNAMIC INTERCHANGE BETWEEN FLORISTIC COMPOSITION AND INDUSTRIAL POLLUTION: AN ECOLOGICAL PERSPECTIVE

UJALA ELAZ<sup>1</sup>, SIFUJAL MULK KHAN<sup>2,3,\*</sup>, ABDULLAH<sup>4</sup>, SABIA JEHANGIR<sup>5</sup>,  
NOOR HUSSAIN CHANDIO<sup>6</sup>, QURAT-UL-AIN<sup>7</sup>, RABIA AFZA<sup>8</sup>,  
ZARRIN FATIMA RIZVI<sup>9</sup>, AND NOREEN KHALID<sup>10</sup>

<sup>1</sup>Department of Plant Sciences, Quaid-i-Azam University, Islamabad, Pakistan

<sup>2</sup>Pakistan Academy of Sciences Islamabad, Pakistan

<sup>3</sup>International Society of Ethnobotany

<sup>4</sup>Department of Geography, Sd. University, Khairpur, Pakistan

<sup>5</sup>Department of Botany, Hazara University, Mansehra, Pakistan

<sup>6</sup>Department of Botany, Government College Women University, Sialkot, F-1110, Pakistan

<sup>7</sup>Corresponding author I. email: [shujaqar@gmail.com](mailto:shujaqar@gmail.com)

<sup>8</sup>Corresponding author II. email: [noreenbano@gmail.com](mailto:noreenbano@gmail.com)

### Abstract

The distribution of plants over the earth's surface is not even or random but follows a particular geographical pattern. The variation in the floristic composition of plants can be attributed to various environmental factors. The current study was conducted in Sukkur, Pakistan, a prominent industrial hub of the country, to assess the impact of industrial pollution on floristic composition. Vegetation was sampled using standard quantitative ecological techniques. A total of 179 quadrats were established across three sites based on pollution gradient. The soil samples from each quadrat were examined using standard laboratory techniques to assess various physicochemical parameters and the concentrations of heavy metals. The collected data were analyzed using PCORD and CANOCO software. Our findings indicate that this region exhibits a diverse range of plants from twenty-eight phytogeographic regions, highlighting its rich biodiversity. The most dominant phytogeographic elements were Cosmopolitan (13.3%), followed by Tropical (10.7%); in contrast, the least dominant ones were Western Himalayan, Sub-cosmopolitan, Sino-Japanese, sub-Himalayan, Indo-Chinese, etc., each represented by one member. Poaceae was the dominant family represented by 16 species (3.7%), followed by Asteraceae 18 (3.7%). We observed that floristic diversity decreased as we moved from a less polluted area to a more polluted area. In addition, local residents of the region dispose of cow dung and other household waste along the study region. This adds organic matter and heavy pollutants, coupled with industrial waste, to the environment and has a crucial impact on the distribution of phytogeographic elements in this region. Therefore, we believe industrial pollution has a remarkable role in the distribution of phytogeographic elements. It is suggested that semi-Tropical and Tropical elements distributed in this study should be protected because of their narrow geographic range.

**Key words:** Phytogeography, Floristic regions, Phyto-diversity, Industrial pollution, Cosmopolitan and Tropical

### Introduction

Phytogeography is the study of plant spatial interactions in the present and past (Wickens, 2008), focusing on explaining the range of plants in terms of origin, dispersal, and evolution (Vukleva *et al.*, 2021). Plants and geographical areas are dispersed in a particular pattern on the surface of the earth as a result of numerous environmental variables (Zeb *et al.*, 2021), such as habitat characteristics, hydrology, topography, soil types, and climate change (Bashir *et al.*, 2010; Munir *et al.*, 2020; Pearson & Dawson, 2003). Environmental factors are the main drivers that govern species distribution directly if their values change beyond the eco-physiological tolerance level of the species of an area (Thammant *et al.*, 2021; Wittinger & Peinkovcova, 2021). It is necessary to understand the extent to which these factors play a role in the floristic composition of an area. Industrial pollution is one of the most critical factors in determining the vegetation of an area. It is argued that changes in the environmental conditions may cause the extinction of some species in the subtropical region, which will significantly affect the diversity and, hence, the ecosystem (Deb *et al.*, 2018). To conserve biodiversity, one should better understand an area's phytogeographic distribution and floristic inventory of vegetation (Harris *et al.*, 2012; Jehangir *et al.*, 2024; Qian, 2001; Zankovskii *et al.*

*et al.*, 2024). The impact of natural vegetation on geographic locales extends significantly beyond the influence of climate, geology, and soil typologies as it increasingly grapples with the pervasive effects of anthropogenic activities, notably environmental pollution (Chaturvedi, 2014; Ejaz *et al.*, 2024). This pollution, arising from diverse sources such as industrial emissions, waste disposal practices, and the release of effluents, catalyzes profound and rapid transformations within natural ecosystems (Bayer *et al.*, 2021). These transformations are characterized by marked changes in ecosystem structure, functionality, and species composition, thereby establishing pollution as a dominant and disruptive force within the natural environment (Haq *et al.*, 2020). The detrimental impacts of pollution are particularly evident in industrial activities, where the improper disposal of waste materials, especially those associated with mineral extraction processes, triggers a cascade of environmental repercussions. As elucidated by (Banerjee *et al.*, 2019), industrial pollution exerts a direct and far-reaching impact on natural vegetation, precipitating rapid alterations in the ecological balance. This scenario accentuates the critical need for developing, implementing, and rigorously enforcing environmental regulations and pollution control mechanisms, emphasizing international collaboration and compliance (Ahriz *et al.*, 2018).



## **Harnessing Nature's Remedies: Assessing Phytoremediation Potential and Physiological Responses of Indicator Plant Species**

### **5.1 Introduction**

In the modern era of globalization and swift industrial expansion, pollution has emerged as a critical concern, carrying far-reaching implications. The relentless proliferation of industries globally is a key contributor to this escalating pollution crisis, underlining the urgency for immediate and efficacious interventions. Among the most pressing environmental challenges is the management of wastewater (Appannagari, 2017). Annually, an estimated 359 billion cubic meters of wastewater are generated worldwide; almost half of this wastewater, replete with various contaminants, is discharged into the natural environment without any form of treatment. This situation not only poses a direct threat to ecosystems but also highlights a significant gap in environmental management practices that needs to be addressed (Jones, van Vliet, Qadir, & Bierkens, 2021; Ukaogo, Ewuzie, & Onwuka, 2020). Polluted water from urban and industrial sources releases harmful elements and heavy metals into ecosystems, degrading the environment, harming aquatic systems, and disrupting soil vital for plant growth (Vodyanitskii, 2013). Addressing heavy metal pollution is paramount due to its extensive negative impact on biological entities and human beings. Different methods, including physical, chemical, and biological interventions, have been used for environmental remediation. However, each approach has limitations (Chai et al., 2021). Chemical and physical methods, though effective, are labor-intensive and costly and can disrupt soil microbial ecosystems, potentially leading to secondary pollution (Ayangbenro, Babalola, & health, 2017; Dixit et al., 2015; Naeem et al., 2021). Achieving lasting environmental resilience needs focused research to devise effective and ecologically sustainable strategies, particularly in the context of heavy metal pollution.

Interestingly, despite the harsh and toxic conditions such pollutants create, certain organisms demonstrate remarkable resilience, notably some plant species. Their ability to thrive in these environments offers valuable insights into potential natural mechanisms of pollution mitigation. Understanding and harnessing these natural adaptations could pave the way for innovative, biologically-based solutions to counteract the adverse effects of heavy metal contamination, aligning with the broader goal of sustainable environmental management. (Kanwar, Sharma, Srivastav, Rani, & Research, 2020; Levin et al., 2001).



Organisms capable of surviving in heavy metal-laden environments can serve as vital indicators of pollution in specific habitats or biomes (Freeman et al., 2019). Certain plant species within these groups exhibit an extraordinary capacity to absorb or degrade heavy metals at concentrations significantly higher than those in typical environments. These species effectively act as natural sinks, playing a crucial role in the remediation of polluted sites (Nascimento, Biondi, Silva, & Lima, 2021; Samudro & Mangkoedihardjo, 2020). Moreover, these plants have evolved advanced phytoremediation techniques to cope with and counteract the polluted environments they face. This adaptation highlights the considerable potential of these species in contributing to environmental restoration and pollution control initiatives (Del Buono, Terzano, Panfili, & Bartucca, 2020).

### ***5.1.2 Phytoremediation: A Green Approach to Environmental Cleanup***

The term 'phytoremediation' was introduced by (Raskin, Kumar, Dushenkov, & Salt, 1994), combining the Greek 'Phyto' (meaning plants) with the Latin 'remedium' (to restore). This approach is increasingly recognized as an eco-friendly and innovative solution for managing hazardous levels of heavy metals in polluted environments (Oladoye, Olowe, & Asemoloye, 2022; Priya, Muruganandam, Ali, & Kornaros, 2023; Sarfraz et al., 2022b; H. Singh & Pant, 2023). This technique utilizes the natural capacity of plants to absorb heavy metals from their surroundings (Etim, 2012; Muthusaravanan et al., 2018). Diverse plant species demonstrate the ability to bio transform, translocate, and accumulate heavy metals (Ejaz et al., 2022b) utilizing phytoremediation strategies to thrive in challenging environments (Ahmad et al., 2023b; S. M. Khan, Haq, Khalid, Ahmad, & Ejaz, 2022). The phytoremediation process encompasses multiple stages tailored to different types of pollution (Ghosh & Singh, 2005). Inorganic pollution has stages that include phytoextraction and Phyto stabilization (Ghosh & Singh, 2005). Phytoextraction involves certain plant species absorbing and accumulating heavy metals from polluted ecosystems in their roots or shoots (Robinson, Anderson, & Dickinson, 2015). Phyto stabilization, conversely, involves plant root exudates demobilizing and binding heavy metals within the soil matrix, reducing their bioavailability (Bolan, Park, Robinson, Naidu, & Huh, 2011). For organic pollution, the stages are phytodegradation/Phyto transformation, biodegradation, and hemofiltration (Asante-Badu et al., 2020). Phytodegradation or Phyto transformation refers to the breakdown of heavy metals absorbed by plants, either through metabolic processes within the plants or externally through compounds released by plant roots (McCutcheon & Schnoor, 2003). Rhizo-filtration involves using plant roots to purify polluted water by filtering out contaminants, a process effectively



utilized by aquatic species such as duckweed (*Lemna minor* L.), sharp dock (*Polygonum amphibium* L.), and pennywort (*Hydrocotyle umbellata* L.), as noted by (Vara Prasad & de Oliveira Freitas, 2003).

### **5.1.3 Mechanisms of Heavy Metal Uptake by Plants**

Plants are remarkably efficient in absorbing essential and non-essential nutrients from their environment, a trait crucial in phytoremediation (Tangahu et al., 2011). They achieve this by producing chelating agents that induce changes in pH and redox reactions, facilitating the solubilization and accumulation of nutrients from low concentrations and insoluble compounds in soils (Sumiahadi & Acar, 2018). This process extends to the uptake, translocation, and storage of heavy metals, which often chemically resemble essential elements (Nouri et al., 2009). The cell membranes of plants are equipped with various transport mechanisms for ion uptake and translocation. These include channels, proteins that assist in ion transport; proton pumps or ATPases, which create an electrochemical gradient using energy; and co- and anti-transporters, which use the gradient produced by ATPases for active ion uptake (Z. Yang et al., 2022). Each system is specific to various nutrients or ions, and plant uptake and translocation processes are meticulously regulated. Plants generally avoid accumulating trace elements beyond their metabolic needs, typically 10-15 ppm (S. S. Kumar, Kadier, Malyan, Ahmad, & Bishnoi, 2017). However, hyperaccumulator species are a notable exception, capable of absorbing heavy metals in concentrations reaching thousands of ppm (X. Yang, Feng, He, & Stoffella, 2005). Hyperaccumulators have adapted mechanisms to avoid metal toxicity, such as storing heavy metals in vacuoles (Sarma, 2011). The process of transpiration also contributes to this mechanism, where evaporation from leaves aids in drawing nutrients into plant roots and consequently transferring contaminants into plant shoots (Usman, Al-Ghouti, & Abu-Dieyeh, 2018). Hyperaccumulators, often used in phytoextraction, require little maintenance to remove toxic elements from their environment, as observed in studies (Sumiahadi & Acar, 2018). These phytoremediators can concentrate or uptake heavy metals at levels 100-1000 times greater than non-accumulator or excluder species, demonstrating their effectiveness in cleansing polluted environments (Chehregani, Noori, & Yazdi, 2009). To mitigate the risk of heavy metal toxicity, phytoremediators utilize various mechanisms, including producing compatible osmolytes like proline, glycerol, glycine betaine, and sorbitol (Ahmad et al., 2023a). These substances enhance their ability to thrive in contaminated environments (Bhuiyan et al., 2019; Borhannuddin Bhuiyan et al., 2019; Ejaz et al., 2023b). The accumulation and synthesis of



these osmolytes, recognized as Osmo protectants, serve as the initial line of defence against potential stressors imposed by toxic elements, adeptly countering the osmotic stress induced by heavy metals (Arif et al., 2019; Kushwaha, Rani, Kumar, & Gautam, 2015). Among these, proline, a pivotal osmolyte, plays a crucial role in various biotic and abiotic stressors, synthesized through the ornithine/glutamate pathway (Siripornadulsil, Traina, Verma, & Sayre, 2002). Glutamate facilitates proline biosynthesis from glutamic acid, occurring within the cytoplasm or chloroplasts via pyrroline-5-carboxylate, a vital pathway for osmotic stress adaptation (Aslam et al., 2017). Pyrroline-5-carboxylate triggers the generation of reactive oxygen species, subsequently countered through an array of enzymatic and non-enzymatic detoxification mechanisms (Siddique, Kandpal, Kumar, & Microbiology, 2018). Additionally, the quantification of chlorophyll content provides a valuable metric for assessing the ecological implications of pollution on flora within contaminated habitats (Wani et al., 2018). Monitoring chlorophyll levels can reveal the extent of pollution-induced stress on plants, shedding light on the overall ecological health of contaminated ecosystems (Piotto et al., 2018).

#### ***5.1.4 Factors Affecting Heavy Metal Accumulation in Plants***

The accumulation of heavy metals in plants is influenced by various factors, encompassing the type of plant species, the bioavailability of metals, characteristics of the growth medium or soil, root zone properties, vegetative uptake mechanisms, and the introduction of chelating agents, among others (Máthé-Gáspár & Anton, 2005). Identifying plant species with superior remediation capabilities is a pivotal aspect of this process (Nouri et al., 2009). Plant species' distinctive attributes significantly impact heavy metal accumulation (Nouri et al., 2009). Diverse strategies have been developed to modify soil properties, such as altering pH levels, applying fertilizers, and introducing chelators, to enhance their remediation potential (Dipu, Kumar, & Thanga, 2012). For instance, changes in soil pH, organic matter content, and phosphorus concentration can affect the accumulation of lead in plants (F. Zeng et al., 2011). Furthermore, the rhizosphere is pivotal in phytoremediation (Agarwal, Giri, & Rani, 2020). It is a zone where heavy metals can be absorbed or metabolized within plant tissues. Morphological adaptations, such as an increase in root diameter and a reduction in root elongation, indicate plant species' responses to polluted soil conditions (Brandão et al., 2018). Environmental conditions also exert a significant influence on the accumulation of heavy metals. Temperature, for instance, affects plant growth and subsequently impacts root length (Cataldo & Wildung, 1978). A



comprehensive understanding of contaminants' mass balance and metabolic fate is essential to demonstrate the efficacy of phytoremediation (Ejaz et al., 2022a). Moreover, the accumulation of heavy metals is contingent on factors like bioavailability in the water phase, retention time, and interactions with other elements. When heavy metals amalgamate with the soil, the redox potential, pH levels, and organic matter concentration become pivotal in determining whether metals exist in an ionic form or one that is accessible to plants (Brezonik, King, & Mach, 2020). The plant species can influence the soil by modifying its pH and oxygenation, which affect metal availability. The addition of synthetic chelating agents represents a viable approach to enhance phytoremediation capabilities (Yaashikaa, Kumar, Jeevanantham, & Saravanan, 2022). These agents facilitate the leaching of heavy metals into the soil. It is worth noting that the bioavailability of heavy metals decreases in soils with a pH ranging from 5 to 6; therefore, chelating agents may be employed in alkaline soil remediation efforts. For instance, exposure of plant species to Ethylenediaminetetraacetic acid (EDTA) for a duration exceeding two weeks has been demonstrated to improve the translocation of metals, with significant results achieved at an EDTA application rate of 5 mmol/kg (Saleem et al., 2020). Furthermore, the release of organic acids like oxalate and citrate by plant roots also exerts an influence on the availability of heavy metals. The presence of ligands can further modulate the accumulation of heavy metals by forming complexes with the metals, ultimately influencing the leaching of contaminants beneath the rhizosphere (Diarra, Kotra, & Prasad, 2021).

#### ***5.1.5 Advantages of Phytoremediation***

Phytoremediation, a relatively novel technology, offers many advantages for addressing contamination in water, soil, and the surrounding air (Adeoye, Adebayo, Afodun, & Ajijolakewu, 2022). Notably, it is a more publicly accepted and environmentally friendly alternative than traditional chemical and physical remediation methods. This approach to contaminant reduction is cost-effective and remarkably well-suited for remediating extensive areas plagued by hazardous elements and pollution (Nong et al., 2023). One of its most compelling attributes is its cost-effectiveness, with costs typically running 60-80% lower than those associated with conventional technologies (Nong et al., 2023). Consequently, phytoremediation eliminates the need for highly specialized personnel and costly equipment, making it an economically viable choice. It proves particularly effective in large areas with moderately polluted surface soils, offering a sustainable solution (Kanwar, Sharma, Srivastav, & Rani, 2020). Phytoremediation's versatility is also worth highlighting, as it can effectively



target a broad spectrum of environmental contaminants, whether organic or inorganic (Garg & Roy, 2022). This technology extends its reach to encompass radionuclides and toxic elements while minimizing environmental disruption. Regarding in-situ applications, phytoextraction, a critical method within phytoremediation, significantly degrades contaminants directly within the soil, resulting in minimal soil disturbance (Shen et al., 2022). This approach leaves topsoil usable, making it suitable for future agricultural purposes. Additionally, phytoremediation stands as an eco-friendly alternative to more invasive and damaging remediation approaches, such as thermal vaporization, incineration, solvent-based treatments, and soil-washing techniques (Saxena, Purchase, Mulla, Saratale, & Bharagava, 2020). Unlike these methods, phytoremediation preserves the biological components of the soil and maintains its physical and chemical features, ultimately producing a more viable waste product. It is worth noting that phytoremediation is considered one of the most significant ecological cleanup technologies for remediating contaminated soil, often referred to as "green technology" or the "green liver of the earth." (Ogundele & Anaun, 2022). Another advantage of phytoremediation lies in its ability to generate recyclable metal-rich plant residues, further contributing to sustainability (Amabogha, 2022). Moreover, the biomass generated during phytoremediation procedures can be economically valorized through bioenergy, offering a sustainable energy source (Saxena et al., 2020). In cases where phytoremediation is applied using oil crops, the resulting plant oil can be utilized for biodiesel production, providing yet another avenue for bioenergy generation. On a larger scale, the stored potential energy within harvested biomass can be harnessed to generate thermal energy, underscoring the multifaceted benefits of phytoremediation as a versatile and eco-conscious remediation solution (Nissim & Labrecque, 2021).

#### ***5.1.6 Challenges in Phytoremediation***

While offering several advantages, phytoremediation is accompanied by a range of limitations that warrant consideration. Firstly, it is a time-consuming process, often spanning multiple growing seasons, which may not be suitable for sites with urgent or severe contamination requiring immediate attention (Shen et al., 2022). Numerous factors can hinder the efficacy of phytoremediation, including soil characteristics, root length, plant growth periods, environmental conditions, the magnitude of contaminants, and the complex interplay between soil and contaminant chemistry (Lee et al., 2021). One critical limitation is ensuring contact between plant roots and contaminants, which can only occur if roots extend deep enough or if contaminants are close to the root zone (Lee et al., 2021). Root age is another



crucial factor influencing phytoremediation efficiency, with younger plants displaying more robust physiological activity and greater remediation potential than their older counterparts (Anerao, Kaware, kumar Khedikar, Kumar, & Singh, 2022). The inhibitory effects of elevated levels of heavy metals on plant growth pose yet another constraint, necessitating the use of specific hyperaccumulator plant species capable of tolerating and accumulating these toxic elements at high concentrations (Farooqi, Hussain, Ayub, Qadir, & Ilic, 2022). Environmental factors further shape the success of phytoremediation efforts, highlighting the need for site-specific considerations. Additionally, the harvested plants used in phytoremediation must be promptly and properly disposed of to prevent further contamination (Gavrilescu, 2022). Nevertheless, there remains a risk that toxic elements could enter the food chain through animals and insects, underscoring the need for caution in managing the aftermath of phytoremediation projects (Pouresmaieli, Ataei, Forouzandeh, Azizollahi, & Mahmoudifard, 2022).

#### ***5.1.7 Physiological Responses of Plants to Heavy Metal Pollution***

In the face of environmental pollution, diverse plant species exhibit remarkable physiological responses to ensure their survival and resilience (P. Zeng et al., 2020). These responses often involve the production of compatible osmolytes, such as glycine betaine, glycerol, sorbitol, and proline, which serve as crucial defense mechanisms (Thakur et al., 2022). These osmolyte molecules are typically water-soluble at specific physiological pH levels and tend to accumulate within the cell's cytosol (Sadiq, Zaid, & Khan, 2020). Their roles extend to regulating cellular osmotic pressures, detoxifying Reactive Oxygen Species (ROS), and ensuring the stability of proteins and membrane integrity (Hoque et al., 2021). Proline, in particular, emerges as a pivotal player in coping with various abiotic and biotic environmental stresses, including drought, pathogen attacks, salinity, nutrient deficiencies, heavy metal accumulation, pollution, and temperature fluctuations (Parmar, Dave, Sudhir, Panchal, & Subramanian, 2013). This amino acid is crucial for primary metabolism and is accumulated by various plant species when exposed to environmental stressors, including lithospheric pollution. The proline level within plant cells is finely regulated through processes involving transport between cells, biosynthesis, and catabolism (Saradhi, 1991). Proline biosynthesis occurs through two distinct pathways: the glutamate and ornithine pathways. The glutamate pathway involves proline synthesis from glutamic acid, leading to the formation of Pyrroline-5-Carboxylate in the chloroplast or cytoplasm. This pathway assumes significance in osmotic stress, nitrogen deficiency, and other stressful physiological



conditions (Pandian et al., 2020). However, it is important to note that excessive Pyrroline-5-Carboxylate production can lead to the generation of ROS, initiation of apoptosis, and damage to cellular components (Thakur et al., 2022). Consequently, plants tend to degrade it swiftly when the stressors are alleviated, as ROS can induce programmed cell death (Adejumo, Awoyemi, & Togun, 2020). ROS also inflict damage by reducing chlorophyll content, affecting membrane fluidity, and inducing lipid peroxidation (Mulenga, Clarke, & Meincken, 2020). In response, plants have evolved various enzymatic and non-enzymatic detoxification mechanisms to counteract ROS (Pandian et al., 2020). The extent of oxidative damage, often measured as lipid peroxidation in malondialdehyde, indicates environmental stress (Gill, Kanwar, Rodrigues dos Reis, & Ali, 2022). Additionally, the determination of physiological traits, such as chlorophyll concentration, offers insights into the impact of environmental pollution on plant species. It serves as a tool for assessing the interference of pollution or stress with photosynthesis, as reduced chlorophyll levels signify negative consequences (Adejumo et al., 2020). Chlorophyll concentration is regarded as an adaptive trait in plant species thriving in polluted or stress-prone habitats. The rhizosphere, the region surrounding plant roots, holds particular significance in phytoremediation efforts. Microbial populations in the rhizosphere are notably higher than in root-free zones and have direct associations with plant species. These microorganisms play a vital role in mobilizing potentially toxic element (PTE) ions, thereby increasing their bioavailability for remediation purposes (Y. Wang et al., 2022).

Industrial pollution directly or indirectly influences nearly all organisms (Behera & Reddy, 2002). However, certain resilient and indicative plant species have evolved robust defense mechanisms to thrive in inhospitable habitats. They employ strategies akin to phytoremediation while adapting physiologically to confront these challenging circumstances (Galiulin, Bashkin, Galiulina, & Birch, 2001). This chapter explores the phytoremediation capabilities exhibited by naturally occurring indicator plant species, as previously discussed. It also scrutinizes their physiological responses, explicitly focusing on proline accumulation and reductions in chlorophyll content within industrial waste-polluted ecosystems. The findings and insights from this study apply to this particular polluted ecosystem and can serve as a blueprint for research in diverse pollution-impacted settings. Additionally, we posit that cultivating and propagating naturally occurring plants with proven ecological and physiological benefits can be a viable approach for reforestation efforts, contributing to environmental restoration on a broader scale. Our study stands out for its comprehensive



approach, focusing on how pollution from industrially polluted water affects soil ecosystems and plant health. This study aims to: 1. Assess soil contamination levels in ecosystems influenced by industrially polluted water. 2. Investigate the phytoremediation potential of native indicator plant species in these impacted environments. 3. Examine the physiological responses of these plants to heavy metal stress, changes explicitly in proline synthesis and variations in chlorophyll-a, chlorophyll-b, and total carotenoid contents. This research has the potential to be applied across diverse polluted ecosystems, contributing to environmental sustainability in the face of rapid industrial growth and climate change. Both governmental and non-governmental organizations can utilize these native plant indicators in contaminated areas to rehabilitate the environment and reduce the presence of hazardous substances.

## **5.2 Material and Methodology**

As discussed in the preceding chapter, we identified indicator plant species using indicator species analysis. These identifications were supported through canonical correspondence analysis and structural equation modeling. To assess the phytoremediation potential of the indicator plants within the highly polluted zone, we collected these plants from all three pollution gradient zones: low, moderate, and high. This comprehensive approach allowed us to compare and analyze their phytoremediation capabilities and physiological responses to soil heavy metal pollution in the Aik-stream region. Before the analysis, we meticulously separated plant roots and shoots, ensuring they were free from any surface debris, soil, or other contaminants. Subsequently, the plant specimens were subjected to drying in an Oven (DSO-300D) and ground into a fine powder.

### ***5.2.1 Quantification of heavy metals***

An Atomic Absorption Spectrophotometer was employed to measure the concentrations of eight heavy metals, including copper (Cu), zinc (Zn), chromium (Cr), lead (Pb), cadmium (Cd), nickel (Ni), arsenic (As), and mercury (Hg). One gram of the respective material was placed in a 250 mL conical flask for each indicator plant's root, shoot, and soil. To facilitate digestion, a 3:1 mixture of nitric acid and perchloric acid (10 mL) was added to each flask and left to react overnight, totaling 24 hours. Subsequently, all samples were digested on a hot plate at 150 degrees Celsius until the solution reached a pale yellowish colour. After cooling, 40 mL of distilled water was added, and the samples were filtered using Whatman No. 42 filter paper. Heavy metal concentrations were quantified using the VARIAN AA240FS Atomic Absorption Spectrophotometer. The final metal concentrations were determined using the formula (Eq. 5.1).



$$\text{Heavy metal concentration (mg / kg)} = \frac{\text{AA reading} - \text{Blank reagents}}{\text{weight (kg)} \times \text{volume raised} \times \text{df}} \dots\dots\dots (5.1)$$

Where: AA = atomic absorption reading, df = dilution factor.

### 5.2.2 Phytoremediation examination

The potential of each indicator plant species of the polluted zone was assessed for phytoremediation using standard metrics for accumulation, transfer, and concentration. Specifically, we employed the Bioaccumulation Coefficient (BAC) (Eq. 5.2), Translocation Factor (TF) (Eq. 5.3), and Biological Concentration Factor (BCF) (Eq. 5.4) as detailed in the study by (Malik, Husain, et al., 2010). A plant is categorized as a Phyto stabilizer when its BCF value surpasses one while its TF value remains below one. The plant is identified as a phytoextractor when the BCF and TF values exceed one. Furthermore, a plant is classified as a hyperaccumulator when its BCF, TF, and BAC values exceed one (Buscaroli, 2017).

$$BCF = \frac{\text{Metal concentration in roots}}{\text{Metal concentration in soil}} \dots\dots\dots (5.2)$$

$$TF = \frac{\text{Metal concentration in shoots}}{\text{Metal concentration in roots}} \dots\dots\dots (5.3)$$

$$BAC = \frac{\text{Metal concentration in shoots}}{\text{Metal concentration in soil}} \dots\dots\dots (5.4)$$

### 5.2.3 Determination of photosynthetic pigments

The photosynthetic components, including chlorophyll-a, chlorophyll-b, and total carotenoids, were assessed for each indicator plant across regions with varying pollution levels, including high, low, and moderate pollution zones. Fresh leaf samples weighing 0.5 grams were first ground using a mortar and pestle in a 4 mL Dimethyl sulfoxide (DMSO) solution to carry out this analysis. Subsequently, the extract was filtered through Whatman No. 1 filter paper to obtain a clear supernatant. These samples were placed in Falcon tubes and heated at 65°C for 4 hours. The concentration of photosynthetic pigments was determined by measuring the absorbance at Optical Density 663 nm and 645 nm using a spectrophotometer instrument, following the method established by (Hiscox & Israelstam, 1979). The standard formulas by (Arnon, 1949), denoted as Eq. 5.6 for chlorophyll-a, Eq. 5.6 for chlorophyll-b and Eq. vii for carotenoids were employed to determine these pigments quantitatively; Z represents the calculated values of pigments.

$$\text{Chlorophyll a} \left( \frac{\text{mg}}{\text{g}} \right) = [1.07 (\text{optical density } 663) - 0.09 (\text{optical density } 645)] = Z \dots\dots\dots (5.5)$$



$$\text{Chlorophyll } b \text{ (mg/g)} = [1.77 (\text{optical density } 645) - 0.280 (\text{optical density } 663)] \dots\dots\dots (5.6)$$

$$\text{Carotenoids content (mg/g)} = (Z \times 4) \dots\dots\dots (5.7)$$

#### 5.2.4 Proline analysis

Proline osmolyte levels were quantified in indicator plants identified from areas with varying pollution levels, including high, low, and moderately polluted zones. This analysis used the methods outlined (Bates, Waldren, & Teare, 1973). A fresh leaf sample weighing 0.5 grams was initially taken and placed in 3% sulphosalicylic acid. The sample was then ground and filtered. Next, a 2 mL portion of the resulting sample was mixed with equal amounts of Ninhydrin and Glacial acid (2 mL each) and heated for 1 hour using a water bath set at 100°C. Subsequently, the samples were allowed to cool to room temperature and maintained below 0°C to halt any ongoing chemical reactions. Toluene (4 mL) was introduced into each sample and left for 50 minutes. Finally, the upper layer of toluene was removed, and absorbance was measured at 520 nm. The quantification of proline accumulation for each sample was determined using a standard available curve, as indicated by Equation 5.8.

$$\text{Proline } (\mu\text{g/g}) = \frac{K \text{ value} \times \text{dilute.Factor} \times \text{optical density } 520\text{nm}}{\text{weight of sample}} \dots\dots (5.8)$$

Where, K = 17.52, dilution factor = 2 and weight of sample = 0.5g

#### 5.2.5 Statistical analysis

##### 5.2.5.1 Ordinary Least Square (OLS)

This study utilizes the model, a standard tool in linear regression analysis, to explore the connection between phytoremediation capacity and the levels of proline, chlorophyll-a, chlorophyll-b, and carotenoids in indicator plants. The OLS model is selected for its proficiency in delineating linear relationships and its ability to estimate the impact of independent variables on a dependent variable. The major purpose is to determine how phytoremediation affects the amounts of proline and other photosynthetic pigments in plants. The following formula gives the OLS model used in this study:

$$Y_i = \alpha + \beta_{\text{proline}} + \beta_1 \text{Chlorophyll}_a + \beta_2 \text{Chlorophyll}_b + \beta_3 \text{Carotenoids} + \epsilon_{ij} \dots\dots\dots (5.9)$$

This approach is helpful because it simply interprets variable values and focuses on reducing the sum of squared errors.



#### 5.2.5.2 Probit Model

The Probit model is used to analyze the possibility of phytoremediation in plants utilizing parameters such as proline, chlorophyll-a, chlorophyll-b, and carotenoids. The model is expressed as 5.10.

$$P(Y = 1)|proline, chlorophyll_a, chlorophyll_b, carotenoids) = \delta(\alpha + \beta_{proline} + \beta_1 Chlorophyll_a + \beta_2 Chlorophyll_b + \beta_3 Carotenoids + \epsilon_{ij} \dots \dots \dots (5.10)$$

This formula determines the probability of phytoremediation activity depending on the explanatory variables supplied, providing a probabilistic explanation of how these components contribute to the phytoremediation procedure in plants.

#### 5.2.5.3 Generalized Mixed Effect Model (GLMM)

The Generalized Mixed Effect Model (GLMM) was used to investigate the effects of fixed components (proline, chlorophyll-a, chlorophyll-b, carotenoids) and various pollution levels (low, moderate, and high) on phytoremediation. Unlike OLS and Probit models, which only address fixed variables, GLMMs effectively integrate Generalized Linear Models (GLMs) with Mixed Effects Models to handle fixed and random variables. This approach evaluates the effects of these variables on phytoremediation outcomes (BCF, TF, BAC) in response to different heavy metals (Cr, Ni, Cu, Cd, Zn, Pb, Hg and As). The model is formulated as 5.11.

$$Y_{ij} = \alpha + \beta_{proline} + \beta_1 Chlorophyll_a + \beta_2 Chlorophyll_b + \beta_3 Carotenoids + b_{zone} + \mu_{ij} \dots \dots (5.11)$$

Equation (iii) indicates a generalized mixed-effect model. Y is phytoremediation, the subset of (BCF, TF and BAC), respectively, in various polluted zones. Whereas  $\alpha$  is the intercept of the generalized mixed effect model, while ( $\beta, \beta_1$  and  $\beta_3$ ) are the coefficients of the fixed variables of the model. Moreover, b indicates the coefficient of the random component model the effect of variation in different zones. Whereas  $\mu_{ij}$  is the error term of the model.



## 5.3 Results

### 5.3.1 Vegetation along the polluted stream

The Indicator Species Analysis (ISA) identified 17 species as indicative of the highly polluted zone, as identified in Chapter 3. These indicator species were collected in three sets of replicates from the three zones along the Aik-Stream. This collection was undertaken for comparative analysis to evaluate their capability for phytoremediation and investigate their physiological reactions to heavy metal contamination, as outlined in Table 5.1.

Table 5.1 List of indicator plants assessed for phytoremediation capacity within the industrial polluted zones of the Aik-Stream in Sialkot, Pakistan.

S. No.	Botanical name	Family	Native/Invasive Status	Growth Form	IVI		
					HPZ	MPZ	LPZ
1	<i>Achyranthes aspera</i>	Amaranthaceae	Native	Herbaceous	242.412	77.2997	43.0656
2	<i>Arundo donax</i>	Poaceae	Invasive	Herbaceous	939.057	384.313	410.531
3	<i>Calotropis procera</i>	Apocynaceae	Native	Shrubby	196.273	126.451	228.856
4	<i>Cannabis sativa</i>	Cannabaceae	Native	Herbaceous	1977.25	860.251	1362.78
5	<i>Chenopodium album</i>	Amaranthaceae	Native	Herbaceous	273.556	29.0046	76.0679
6	<i>Coronopus didymus</i>	Brassicaceae	Invasive	Herbaceous	374.529	191.357	72.3332
7	<i>Cynodon dactylon</i>	Poaceae	Native	Herbaceous	336.082	145.166	281.196
8	<i>Dysphania ambrosioides</i>	Amaranthaceae	Invasive	Herbaceous	298.504	37.1634	101.028
9.	<i>Eclipta alba</i>	Asteraceae	Native	Herbaceous	197.283	79.7221	398.994
10	<i>Ficus carica</i>	Moraceae	Native	oody	420.11	21.25	118.672
11	<i>Koeleria macrantha</i>	Poaceae	Native	Herbaceous	137.947	41.8645	135.649
12	<i>Malva neglecta</i>	Malvaceae	Native	Herbaceous	299.041	174.563	51.8692
13	<i>Ricinus communis</i>	Euphorbiaceae	Invasive	Herbaceous	215.785	37.3869	41.2037
14	<i>Parthenium hysterophorus</i>	Asteraceae	Invasive	Herbaceous	636.721	337.361	336.344
15	<i>Persicaria glabra</i>	Polygonaceae	Native	Herbaceous	95.5179	166.663	203.98



16	<i>Saccharum bengalensis</i>	Poaceae	Native	Herbaceous	205.795	167.385	1087.37
17	<i>Ziziphus nummularia</i>	Rhamnaceae	Native	oody	313.495	192.524	173.867

### 5.3.2 Soil of the Polluted Ecosystem

Soil of the polluted zone had copper concentrations ranging from 5.34-65.12 (average 24.53) mg/kg, Zinc 10.15-81.51 (37.58) mg/kg, chromium from 19.54-266.06 (79.33) mg/kg, lead 2.62-86.65 (32.73) mg/kg, cadmium 0.021-4.867 (1.36) mg/kg, nickel ranged from 0.34-33.62 (8.92) mg/kg. Chromium's highest concentration was reached based on the average metal concentration, followed by Zn > Pb > Cu > Ni > Cd. Whereas soil pH deviated from 6.55-8.75 (an average 7.44), electrical conductivity ranged 1.1-17.9 (4.51) dsm<sup>-1</sup>, total dissolved solids various from 51-757 (281.5), organic matter ambit from 0.05-0.99 (0.53), potassium 243-744 (488.3) mg/kg, phosphorus ambit from 2.5-19.4 (10.45) mg/kg. The soil has a 44.11°C maximum, a 2.84 minimum temperature, and 3.81 mm/day precipitation.

Table 5.2 Summary statistics of soil samples from the study area detailing the sample size [N = 150 (3x)].

Soil Parameters	Mean	St.dev	Minimum	Maximum	Skewness	Kurtosis	Permissible limit	Toxicity
Temperature (°C)	44.1	0.804	2.84	44.1	-12.5	155	26.6	-
Soil moisture	0.630	0.540	0.630	0.630	-12.5	155	-	-
TDS (mg/L)	315	152	51	757	0.646	0.314	500	-
Saturation (%)	34.1	5.40	20.0	58.0	0.953	0.916	-	-
pH	7.54	0.533	6.55	9.20	0.973	0.80	7	-
EC (μS/cm)	4.58	3.79	1.10	17.9	1.94	3.46	110	-
OM (%)	0.39	0.550	0.050	0.990	-0.03	-1.40	0.05	-
P (mg/kg)	10.45	1.71	2.50	19.4	-0.13	0.206	30	-
K (mg/kg)	488.3	78.7	243	744	1.32	0.196	300	-
Cu (mg/kg)	25.7	14.9	5.34	65.1	0.632	-0.55	36	5
Zn (mg/kg)	38.4	17.5	10.2	81.5	0.555	-0.64	50	1
Cr (mg/kg)	80.2	67.8	20.8	266	1.27	0.555	100	2
Pb (mg/kg)	49.8	37.9	2.62	99.8	0.0996	-1.85	85	5
Cd (mg/kg)	1.38	1.22	0.021	4.87	0.983	0.983	0.8	30
As (mg/kg)	0.748	0.678	0.678	1.98	0.598	-1.28	10	0.5
Ni (mg/kg)	20.5	19.8	0.342	67.4	0.464	-1.49	35	5
Hg (mg/kg)	0.179	0.192	0.0080	0.928	0.928	1.98	1	0.27



### 5.3.3 Phytoremediation of heavy metals

A total of 17 indicators of the industrially polluted zone along the Aik-Stream in Sialkot, Pakistan, were *Achyranthes aspera*, *Arundo donax*, *Calotropis procera*, *Cannabis sativa*, *Chenopodium album*, *Coronopus didymus*, *Cynodon dactylon*, *Dysphania ambrosioides*, *Eclipta alba*, *Ficus carica*, *Koeleria macrantha*, *Malva neglecta*, *Ricinus communis*, *Parthenium hysterophorous*, *Persicaria glabra*, *Saccharum bengalensis*, and *Ziziphus nummularia*. The assessment of these species for their ability to remediate heavy metals by examining their accessibility, indicator properties, and Importance Value Index (IVI) information. The study revealed that every species identified as an indicator demonstrated notable phytoremediation potential in managing the observed heavy metal levels. This was determined through an evaluation matrix consisting of the Biological Concentration Factor (BCF), Translocation Factor (TF), and Bioaccumulation Coefficient (BAC). A detailed summary of the statistics can be found in Table 5.3.



Table 5.3 Biological Concentration Factor (BCF), Translocation Factor (TF) and Bioaccumulation Coefficient (BAC) for different heavy metals of indicator species in the highly, moderately, and less polluted zones

Plant Names	Factors	Cr	Ni	Cu	Cd	Zn	Pb	As	Hg
<b>LPZ</b>									
<i>Achyranthes aspera</i>	BCF	0.996	1.0220	0.8312	1.7087	0.3512	0.2928	1.036	0.0110
	TF	0.846	1.0333	0.0466	0.07566	0.4700	0.12681	1.016	0.0030
	BAC	0.843	1.0007	0.0387	0.1292	0.1650	0.0371	0.003	0.0170
<b>MPZ</b>									
<i>Achyranthes aspera</i>	BCF	1.0065	0.83127	1.7087	0.996	1.02595	1.00937	0.008	0.036
	TF	0.7383	0.04661	0.07566	0.846	0.94223	0.98559	0.070	0.034
	BAC	0.6022	0.03875	0.1292	0.843	0.9666	0.9948	0.002	0.056
<b>HPZ</b>									
<i>Achyranthes aspera</i>	BCF	1.0647	1.02595	1.83127	0.85614	1.00937	0.9226	1.001	1.026
	TF	0.9090	0.94223	1.04661	0.8351	0.98559	0.9926	1.040	1.031
	BAC	0.9678	0.9666	1.03875	0.714969	0.9948	0.91588	0.002	0.083
<b>LPZ</b>									
<i>Arundo donax</i>	BCF	1.0006	0.2837	0.42301	2.2276	0.3510	0.2775	0.030	0.014
	TF	0.967	0.0548	0.1406	1.3048	0.69144	0.1379	0.050	0.019
	BAC	0.968	0.0155	0.0594	2.9066	0.2427	0.0383	0.011	0.097
<b>MPZ</b>									
<i>Arundo donax</i>	BCF	1.1792	0.4230	2.2276	1.0006	0.97598	1.01723	0.035	0.054
	TF	0.8758	0.1406	1.3048	0.967	0.9458	0.97818	0.106	0.052
	BAC	0.73837	0.0594	2.9066	0.968	0.9231	0.99504	0.059	0.039
<b>HPZ</b>									
<i>Arundo donax</i>	BCF	1.1593	1.97598	0.4230	0.97524	1.01723	1.00571	0.033	1.091
	TF	0.9421	1.9458	0.1406	0.96408	0.97818	0.98528	0.006	1.081
	BAC	1.0921	0.9231	0.0594	0.9402	0.99504	0.9909	0.004	0.050
<b>LPZ</b>									
<i>Calotropis procera</i>	BCF	1.211	0.058	0.4254	2.4036	0.16681	0.37508	0.050	0.0680
	TF	0.9326	0.21142	0.19767	0.0133	0.2130	0.27038	0.100	0.090
	BAC	1.1295	0.0123	0.0841	0.0133	0.03553	0.1014	0.080	0.035
<b>MPZ</b>									
<i>Calotropis procera</i>	BCF	1.9395	1.4254	2.4036	1.211	0.97917	1.0026	0.031	0.0917
	TF	1.9271	1.1976	0.0133	0.9326	0.94732	1.90792	0.072	0.0732
	BAC	1.8706	0.08410	0.0133	1.1295	1.3413	1.9102	0.084	1.003
<b>HPZ</b>									



<i>Calotropis procera</i>	BCF	0.9367	0.97917	0.4254	0.9745	1.0026	0.97105	1.034	0.045
	TF	0.8980	0.94732	0.1976	0.99784	0.90792	0.9899	1.023	0.078
	BAC	0.8412	1.3413	0.08410	0.97244	0.9102	0.9613	0.010	0.024
<b>LPZ</b>									
<i>Cannabis sativa</i>	BCF	1.0667	0.0358	0.4264	1.4985	0.1178	0.2518	0.064	0.073
	TF	0.930	6.3734	0.3310	1.0496	0.9253	0.9928	0.010	0.057
	BAC	0.9925	0.2282	0.1411	1.5730	0.1090	0.2500	0.011	0.091
<b>MPZ</b>									
<i>Cannabis sativa</i>	BCF	1.1094	0.4264	1.4985	1.0667	0.97584	1.10758	0.064	0.058
	TF	0.82287	0.33104	1.0496	0.930	0.9964	0.8175	0.004	0.064
	BAC	0.44166	0.3310	1.5730	0.9925	0.9302	0.90553	0.010	0.002
<b>HPZ</b>									
<i>Cannabis sativa</i>	BCF	0.9367	0.97584	0.4264	0.99159	1.10758	1.0026	0.064	1.099
	TF	0.8980	0.9964	0.33104	0.9557	1.8175	0.9949	0.014	1.057
	BAC	0.8412	0.9302	0.3310	0.94768	0.90553	0.99754	0.013	0.076
<b>LPZ</b>									
<i>Chenopodium album</i>	BCF	1.029	0.01205	1.4447	0.1542	0.8812	0.1705	0.012	0.042
	TF	0.8874	2.875	0.8993	10.3333	0.9760	1.1090	0.075	0.330
	BAC	0.9137	0.03466	1.2994	1.5934	0.8600	0.1891	0.034	0.034
<b>MPZ</b>									
<i>Chenopodium album</i>	BCF	0.2782	1.14500	0.1542	1.029	0.9507	1.0872	0.050	0.410
	TF	0.5662	0.9613	10.3333	1.8874	0.9717	0.9590	0.003	0.432
	BAC	0.8030	1.1006	1.5934	0.9137	1.7856	1.04275	0.06	0.302
<b>HPZ</b>									
<i>Chenopodium album</i>	BCF	1.1828	0.9507	1.14500	0.81410	1.0872	1.0093	0.054	0.049
	TF	0.8392	0.9717	0.9613	0.9741	0.9590	0.97152	0.034	0.057
	BAC	0.9927	1.7856	1.1006	0.7930	1.04275	0.98063	0.033	0.066
<b>LPZ</b>									
<i>Coronops didymus</i>	BCF	1.202	0.0170	1.1593	2.2684	0.8959	0.94442	0.012	0.042
	TF	1.0545	0.8888	0.9702	0.0706	0.9526	0.9961	0.075	0.330
	BAC	1.2676	0.0151	1.12483	0.1603	0.8535	0.9408	0.034	0.034
<b>MPZ</b>									
<i>Coronops didymus</i>	BCF	0.1416	1.22883	2.2684	1.202	0.94223	1.7131	0.054	0.049
	TF	0.6103	0.98388	0.0706	1.0545	0.9666	1.0377	0.034	0.057
	BAC	0.8030	1.2090	0.1603	1.2676	0.97598	0.74001	0.033	0.066
<b>HPZ</b>									
<i>Coronops didymus</i>	BCF	0.9698	0.94223	1.22883	1.0056	0.7131	1.91271	0.034	0.045
	TF	0.8560	0.9666	0.98388	0.97027	1.0377	1.0096	0.023	0.078



	BAC	0.8302	0.97598	1.2090	0.97571	0.74001	0.9214	0.010	0.024
<b>LPZ</b>									
<i>Cynodon dactylon</i>	BCF	1.081	0.9358	1.36572	1.4231	0.6453	1.08582	0.050	0.410
	TF	0.9266	0.9813	0.8871	1.8282	0.9377	1.94899	0.003	0.432
	BAC	1.0021	0.9183	0.2115	1.3751	0.7515	0.0304	0.06	0.302
<b>MPZ</b>									
<i>Cynodon dactylon</i>	BCF	0.1453	1.4465	2.9629	1.081	0.9458	0.82422	0.012	0.042
	TF	0.6216	1.90180	1.2875	0.9266	0.9231	0.95323	0.075	0.330
	BAC	0.8030	0.30450	1.6296	1.0021	0.97917	0.78568	0.034	0.034
<b>HPZ</b>									
<i>Cynodon dactylon</i>	BCF	1.939	0.9458	1.4465	0.8122	0.82422	1.00487	0.036	0.0110
	TF	1.9130	0.9231	0.90180	0.8585	0.95323	0.99158	0.016	0.0030
	BAC	0.8577	0.97917	1.30450	0.69733	0.78568	0.99641	0.003	0.0170
<b>LPZ</b>									
<i>Dysphania ambrosioides</i>	BCF	1.1186	0.9329	1.2033	0.9381	0.7047	1.0065	0.050	0.0680
	TF	0.9168	0.9663	0.9517	0.9044	0.8400	0.98498	0.100	0.090
	BAC	1.0255	0.901	1.14535	0.84843	0.9033	0.9914	0.080	0.035
<b>MPZ</b>									
<i>Dysphania ambrosioides</i>	BCF	0.2232	1.1493	0.9381	1.1186	1.94732	0.83895	0.050	0.410
	TF	0.5193	0.9869	0.9044	0.9168	1.3413	0.89803	0.003	0.432
	BAC	0.8030	1.13443	0.84843	1.0255	1.97584	0.75341	0.06	0.302
<b>HPZ</b>									
<i>Dysphania ambrosioides</i>	BCF	1.9404	1.94732	1.1493	1.21765	0.83895	1.73116	0.064	0.089
	TF	1.9414	1.3413	0.9869	0.9099	0.89803	0.90227	0.044	0.097
	BAC	0.8853	0.97584	0.13443	0.10801	0.75341	0.6597	0.063	0.096
<b>LPZ</b>									
<i>Eclipta Alba</i>	BCF	1.1277	0.9964	1.00360	1.00420	1.93962	0.9682	0.020	0.045
	TF	1.9022	0.9302	0.96229	1.0181	1.9559	0.83466	0.004	0.032
	BAC	0.0175	0.9507	0.96576	1.02241	0.89827	0.8081	0.066	0.072
<b>MPZ</b>									
<i>Eclipta Alba</i>	BCF	0.2261	1.00360	0.8260	1.1974	0.9964	0.93962	0.022	0.072
	TF	0.6106	0.96229	0.76251	0.9286	0.9302	0.9559	0.045	0.430
	BAC	0.7701	0.96576	0.6298	1.11201	0.9507	0.89827	0.064	0.094
<b>HPZ</b>									
<i>Eclipta Alba</i>	BCF	1.1277	0.9964	1.00360	1.00420	1.93962	0.9682	0.034	0.045
	TF	1.9022	0.9302	0.96229	1.0181	1.9559	0.83466	0.023	0.078
	BAC	0.0175	0.9507	0.96576	1.02241	0.89827	0.8081	0.010	0.024
<b>LPZ</b>									



<i>Ficus carica</i>	BCF	1.008	0.9035	1.2086	1.0172	1.8837	0.9981	0.031	0.0917
	TF	0.8683	0.9125	0.95719	1.0802	1.6046	0.9825	0.072	0.0732
	BAC	0.8755	0.8245	1.15693	1.0989	0.8116	0.9806	0.084	0.003
<b>MPZ</b>									
<i>Ficus carica</i>	BCF	0.0901	1.0821	1.0172	1.008	1.9717	0.89677	0.028	1.073
	TF	0.5886	0.9935	1.0802	0.8683	1.7856	0.94627	0.046	1.437
	BAC	0.9782	1.0750	1.0989	0.8755	0.81762	0.84859	0.065	0.096
<b>HPZ</b>									
<i>Ficus carica</i>	BCF	1.1681	0.9717	1.0821	0.9505	0.89677	0.9001	0.061	1.072
	TF	0.9402	1.7856	0.9935	0.92890	0.94627	0.96987	0.043	1.330
	BAC	0.0983	0.81762	1.0750	0.8829	0.84859	0.8730	0.075	0.064
<b>LPZ</b>									
<i>Koelaria macarantha</i>	BCF	0.9396	0.9527	0.8473	0.9775	1.3024	1.0266	0.072	0.034
	TF	0.8376	0.9722	0.95397	0.9655	0.9594	1.9759	0.430	0.023
	BAC	0.7871	0.9263	0.8083	0.9439	0.79742	0.0019	0.094	0.010
<b>MPZ</b>									
<i>Koelaria macarantha</i>	BCF	0.0851	0.93772	0.9775	0.9396	0.39055	0.9035	0.072	0.034
	TF	0.5993	0.7234	0.9655	0.8376	0.4307	0.76915	0.430	0.023
	BAC	0.0103	0.67841	0.9439	0.7871	0.6533	0.6949	0.094	0.010
<b>HPZ</b>									
<i>Koelaria macarantha</i>	BCF	1.0910	0.39055	0.93772	1.6825	0.9035	1.8053	0.064	0.042
	TF	0.8282	0.4307	0.7234	1.9504	0.76915	1.91059	0.044	0.330
	BAC	0.9036	0.6533	0.67841	0.64872	0.6949	0.7333	0.063	0.034
<b>LPZ</b>									
<i>Malava neglecta</i>	BCF	0.9701	1.0120	1.2405	0.9009	0.7651	1.0184	0.072	0.034
	TF	0.9622	0.9704	0.97633	0.9222	0.9802	0.9806	0.430	0.023
	BAC	0.9335	0.9820	1.2111	0.8309	0.9107	0.99874	0.094	0.010
<b>MPZ</b>									
<i>Malava neglecta</i>	BCF	0.0719	1.23613	0.9009	0.9701	0.51941	0.9295	0.064	0.042
	TF	0.7936	0.7861	0.9222	0.9622	0.75159	0.94745	0.044	0.330
	BAC	0.0112	0.9718	0.8309	0.9335	0.8173	0.8806	0.063	0.034
<b>HPZ</b>									
<i>Malava neglecta</i>	BCF	1.0175	0.51941	1.23613	0.8288	1.9295	0.8976	1.011	0.034
	TF	0.8086	0.75159	0.7861	0.9068	1.94745	0.9146	1.001	0.023
	BAC	0.8228	0.8173	0.9718	0.75159	0.8806	0.8210	0.01	0.010
<b>LPZ</b>									
<i>Ricinus communis</i>	BCF	0.9900	0.9643	0.80924	1.0496	0.83104	1.93177	0.064	0.031
	TF	0.8645	0.75159	0.8983	0.9012	0.91435	1.8798	0.044	0.072



	BAC	0.8559	0.8173	0.72700	0.94605	0.75986	1.81978	0.063	0.084
	<b>MPZ</b>								
<i>Ricinus communis</i>	BCF	0.2669	0.80924	0.98225	1.8420	0.9643	0.83104	0.020	0.042
	TF	0.5929	0.8983	0.99607	1.9311	0.75159	0.91435	0.004	0.330
	BAC	0.0186	0.72700	0.9783	1.7840	0.8173	0.75986	0.066	0.034
	<b>HPZ</b>								
<i>Ricinus communis</i>	BCF	1.9900	1.9643	0.80924	1.0496	0.83104	0.93177	0.064	0.034
	TF	1.8645	1.75159	0.8983	0.9012	0.91435	0.8798	0.044	0.023
	BAC	0.8559	0.8173	0.72700	0.94605	0.75986	0.81978	0.063	0.010
	<b>LPZ</b>								
<i>Parthenium hysterophorous</i>	BCF	1.1379	0.8845	0.95399	0.9549	0.7519	1.0042	0.072	0.042
	TF	0.9892	0.8147	0.9716	0.9361	0.917	0.99452	0.430	0.330
	BAC	0.1257	0.7207	0.9269	0.8939	0.90253	0.9987	0.094	0.034
	<b>MPZ</b>								
<i>Parthenium hysterophorous</i>	BCF	0.2446	0.77186	0.9549	1.1379	0.9643	0.88331	0.064	0.031
	TF	0.4764	0.89346	0.9361	0.9892	0.81762	0.67757	0.044	0.072
	BAC	0.0016	0.68963	0.8939	0.1257	0.39055	0.59851	0.063	0.084
	<b>HPZ</b>								
<i>Parthenium hysterophorous</i>	BCF	1.9422	0.9643	0.77186	1.8631	0.88331	1.16783	1.020	1.042
	TF	1.9005	0.81762	0.89346	1.96033	0.67757	1.78877	1.004	1.330
	BAC	1.8485	0.39055	0.68963	1.82892	0.59851	1.1323	0.066	0.034
	<b>LPZ</b>								
<i>Persiaris glabra</i>	BCF	1.1318	0.9690	0.9781	0.97337	0.8116	1.0088	0.020	0.034
	TF	0.9326	1.0311	0.9292	0.94961	1.3024	0.9876	0.004	0.023
	BAC	1.0555	0.9992	0.9088	0.9243	0.9594	0.9964	0.066	0.010
	<b>MPZ</b>								
<i>Persiaris glabra</i>	BCF	0.1642	0.9213	0.97337	1.1318	0.4307	1.0187	0.064	0.031
	TF	0.5701	0.75050	0.94961	0.9326	0.6533	0.95344	0.044	0.072
	BAC	0.0226	0.6914	0.9243	1.0555	0.51941	0.97131	0.063	0.084
	<b>HPZ</b>								
<i>Persiaris glabra</i>	BCF	1.9881	0.4307	1.9213	0.8034	1.0187	1.23541	0.020	0.034
	TF	1.9144	0.6533	1.75050	0.9644	0.95344	0.73647	0.004	0.023
	BAC	0.9036	0.51941	0.6914	0.77486	0.97131	0.1733	0.066	0.010
	<b>LPZ</b>								
<i>Saccharum bengalensis</i>	0.004	0.072	0.9661	0.9969	0.9815	0.7974	0.93216	0.031	0.031
	0.066	0.084	0.9873	0.97817	0.9793	0.7651	0.9790	0.072	0.072
	BAC	1.1135	0.9539	0.9752	0.9612	0.9802	0.91260	0.084	0.084
	<b>MPZ</b>								



<i>Saccharum bengalensis</i>	BCF	0.1278	0.78681	1.9815	0.985	0.75159	0.98322	0.064	0.042
	TF	0.6627	0.8825	1.9793	1.1300	0.8173	0.9843	0.044	0.330
	BAC	0.0112	0.89147	0.9612	1.1135	0.9643	0.96785	0.063	0.034
<b>HPZ</b>									
<i>Saccharum bengalensis</i>	BCF	1.9294	1.75159	0.78681	1.8965	1.98322	0.81147	0.020	0.034
	TF	0.9130	1.8173	0.8825	0.97989	1.9843	0.80477	0.004	0.023
	BAC	0.8485	0.9643	0.89147	0.8784	0.96785	0.6530	0.066	0.010
<b>LPZ</b>									
<i>Ziziphus nummularia</i>	BCF	1.099	0.9725	1.0075	0.8932	0.9107	1.00085	0.064	0.031
	TF	0.8538	0.97460	0.9720	0.91607	0.8926	0.98922	0.044	0.072
	BAC	0.9386	0.9478	0.9794	0.8183	0.9594	0.9900	0.063	0.084
<b>MPZ</b>									
<i>Ziziphus nummularia</i>	BCF	0.0983	0.7746	0.8932	1.099	0.7519	0.9815	0.072	0.034
	TF	0.5993	0.8357	0.91607	0.8538	0.57306	0.9495	0.430	0.023
	BAC	0.0358	0.6473	0.8183	0.9386	0.7993	0.93195	0.094	0.010
<b>HPZ</b>									
<i>Ziziphus nummularia</i>	BCF	1.1792	1.7519	0.7746	0.9122	1.9815	0.75331	0.064	1.042
	TF	0.7352	1.57306	0.8357	0.9481	1.9495	0.96911	0.044	1.330
	BAC	0.8669	0.7993	0.6473	0.86496	0.93195	0.73005	0.063	0.034



The assessment of various plant species as Phyto-extractors, Phyto-stabilizers, and hyper-accumulators of specific heavy metals within an industrially polluted water ecosystem. The hyper-accumulator indicator plants, including *Achyranthes aspera* hyper-accumulator of (Ni, Cu), *Calotropis procera* (Pb, Cr), *Cannabis sativa* (Cd, Cu), *Ficus carica* (Cd, Cu), *Ricinus communis* (Pb, Cd), and *Parthenium hysterophorous* (Cd, Cr, Pb), were among the top plants out of 17 indicator species due to their maximum capacity for surviving in polluted environments. This insight is critical for developing effective phytoremediation strategies in areas affected by industrial pollution. (Table 5.4).

Table 5.4 The phytoremediation capacity of identified indicator plant species for various heavy metals in the industrially polluted zone/highly polluted zone is shown below, with top-performing hyper-accumulators highlighted in bold.

S. NO	Botanical Names	Phyto-extractor	Phyto-stabilizer	Hyper-accumulator
1	<b><i>Achyranthes aspera</i></b>	Ni, As, Hg	Cd, Cr, Cu, Zn, Pb	Ni, Cu
2	<i>Arundo donax</i>	Ni, Hg	Cr, Pb, Zn	Cd
3	<b><i>Calotropis procera</i></b>	As, Ni	Cd, Cu, Zn	Pb, Cr
4	<b><i>Cannabis sativa</i></b>	Zn, Hg	Cr, Pb	Cd, Cu
5	<i>Chenopodium album</i>	Cd	Cr, Pb	-
6	<i>Coronops didymus</i>	Pb, Ni	Cu, Cd	Cr
7	<i>Cynodon dactylon</i>	Ni, Cr	Cu, Pb	-
8	<i>Dysphania ambrosioides</i>	Cr, Ni	Cu, Cd, Ni, Pb	Zn
9	<i>Eclipta Alba</i>	Cr, Zn	Ni, Cu	Cd
10	<b><i>Ficus carica</i></b>	Zn, Hg	Cr, Cd	Cd, Cu
11	<i>Koelaria macarantha</i>	Cd, Pb	Cr, Zn	-
12	<i>Malava neglecta</i>	Cd, Pb	Cr, Ni, Cu, Pb	-
13	<b><i>Ricinus communis</i></b>	Cr, Ni	Cd	Pb, Cd
14	<b><i>Parthenium hysterophorous</i></b>	As, Hg	Cd, Cr	Cd,Cr,Pb
15	<i>Persiaria glabra</i>	Cr, Cu	Pb, Zn,	-
16	<i>Saccharum bengalensis</i>	Ni, Zn	Cu, Cd, Pb	-
17	<i>Ziziphus nummularia</i>	Cr, Ni, Pb, Hg	Cr, Cu, Pb	-

#### 5.3.4 Plant physiological responses

The study measured chlorophyll-a, chlorophyll-b, and total carotenoids in 17 indicator plant species observed. As heavy metal pollution increased, there was a corresponding decrease in these crucial photosynthetic pigments. A notable reduction in chlorophyll content was evident when comparing plants from low pollution zones (LPZ) to those from high pollution zones (HPZ), as detailed in Figure 5.1 a-c and Table 5.4. Furthermore, the research identified a direct correlation between the proline levels and the severity of heavy metal



pollution. It was observed that plants growing in environments with lower pollution exhibited less accumulation of proline, a compound commonly associated with stress response in plants (Figure 3.1d). Summary statistics of the Bioaccumulation Coefficient (BAC), Translocation Factor (TF), Biological Concentration Factor (BCF), proline, chlorophyll-a, chlorophyll-b, and carotenoids of heavy metals can be observed in Table 5.6.

### 5.3.5 Chlorophyll-a

The highest chlorophyll-a levels were observed in *Koelaria macarantha* at 2.586 mg/g, with *Persiaris glabra* close behind at 2.557 mg/g. These were followed by *Malva neglecta* at 1.584 mg/g, *Malvastrum coromandelianum* at 1.548 mg/g, *Saccharum bengalensis* at 1.640 mg/g, *Ziziphus nummularia* with 1.648 mg/g, and *Arundo donax* at 1.598 mg/g within the Less Polluted Zone (LPZ) of the IPWE. On the other end of the spectrum, the least chlorophyll-a content was observed in *Coronops didymus* at 0.662 mg/g, *Achyranthes aspera* at 0.743 mg/g, *Cynodon dactylon* at 0.767 mg/g, *Eclipta Alba* at 0.760 mg/g, as detailed in Figure 2 and Supplementary Table 2. In the High Pollution Zone (HPZ), *Dysphania ambrosioides* had the highest recorded chlorophyll-a concentration at 0.709 mg/g, followed by *Eclipta Alba* at 0.721 mg/g, *Koelaria macarantha* at 0.678 mg/g, *Cannabis sativa* at 0.609 mg/g, *Ziziphus nummularia* at 0.682 mg/g, and *Achyranthes aspera* at 0.649 mg/g. Conversely, the lowest concentrations in this zone were found in *Chenopodium album* at 0.419 mg/g, *Coronops didymus* at 0.410 mg/g, *Calotropis procera* at 0.449 mg/g, *Arundo donax* at 0.460 mg/g, and *Malvastrum coromandelianum* at 0.508 mg/g, which can also be referred to in Figure 5.1 and Table 5.5.



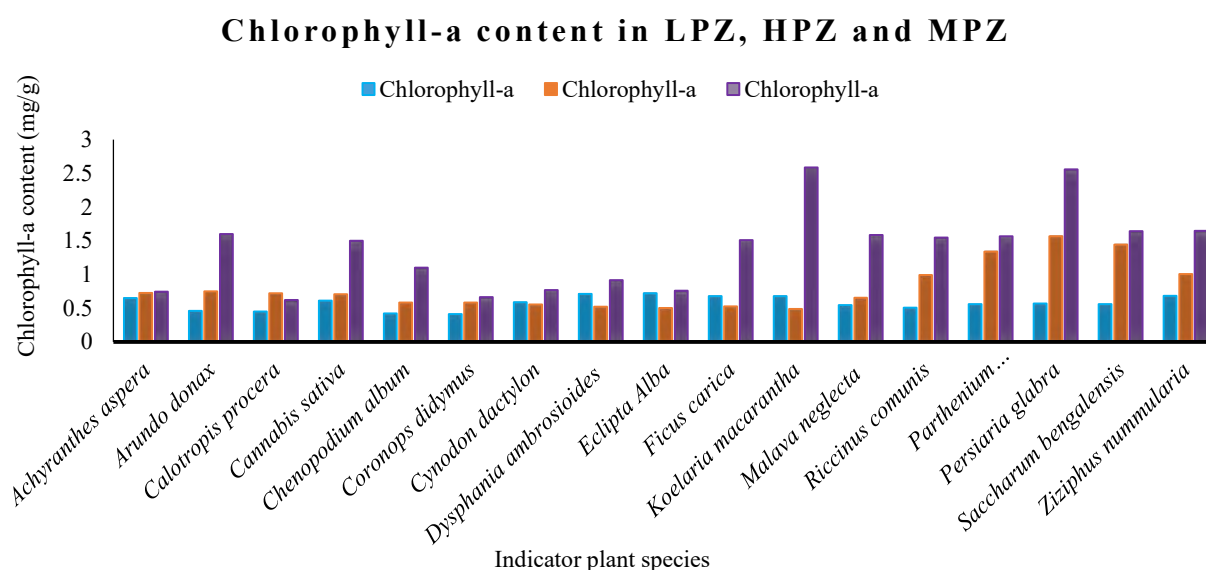


Figure 5.1 Chlorophyll-a content (mg/g) of the identified indicators plant species present in the Less Polluted Zone (LPZ), Moderately Polluted Zone (MPZ) and Highly Polluted Zone (HPZ) of the Industrially Polluted-Water Ecosystem (IPWE).

### 5.3.6 Chlorophyll-b

In the High Polluted Zone (HPZ), *Ficus carica* exhibited the greatest concentration of chlorophyll-b at 0.512 mg/g, followed by *Koelaria macarantha* with 0.4721 mg/g, *Malava neglecta* at 0.4067 mg/g, *Malvastrum coromandelianum* at 0.384 mg/g, and *Ziziphus nummularia* at 0.342 mg/g. The lowest chlorophyll-b levels in HPZ were recorded in *Coronops didymus* at 0.167 mg/g, *Calotropis procera* at 0.171 mg/g, and *Arundo donax* at 0.181 mg/g. Within the Moderate Polluted Zone (MPZ), the highest levels of chlorophyll-b were found in *Persiaria glabra* at 1.4149 mg/g, *Parthenium hysterophorous* at 1.189 mg/g, and *Saccharum bengalensis* at 1.294 mg/g, as indicated in Figure 5.1b and Table 3.4. As for the Less Polluted Zone (LPZ), the maximum chlorophyll-b concentration was noted in *Persiaria glabra* at 2.227 mg/g, followed by *Parthenium hysterophorous* at 2.166 mg/g, *Malvastrum coromandelianum* at 1.8842 mg/g, and *Calotropis procera* at 1.465 mg/g. On the contrary, the minimum amounts were found in *Cynodon dactylon* at 0.349 mg/g, *Cannabis sativa* at 0.450 mg/g, *Achyranthes aspera* at 0.574 mg/g, *Ficus carica* at 0.648 mg/g, and *Eclipta Alba* at 0.735 mg/g. These findings are documented in Figure 5.2 and Table 5.5.



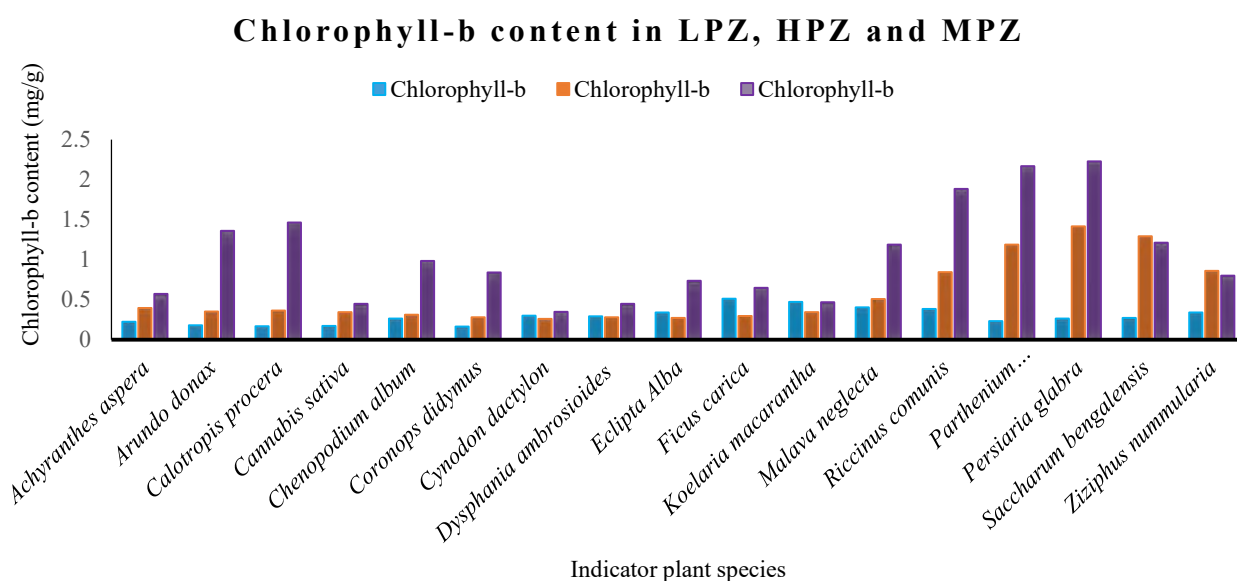


Figure 5. 2 Chlorophyll-b content (mg/g) of the identified indicators plant species present in LPZ, MPZ and HPZ of the IPWE.

### 5.3.7 Carotenoids

In the High Polluted Zone (HPZ), the analysis revealed that *Malvastrum coromandelianum* demonstrated the highest carotenoid concentration, peaking at 3.427 mg/g. It was closely followed by *Persiaria glabra* and *Saccharum bengalensis*, with concentrations of 3.068 mg/g and 3.007 mg/g, respectively. Other significant measurements included *Malava neglecta* at 2.999 mg/g, *Calotropis procera* at 2.813 mg/g, *Parthenium hysterophorous* at 2.787 mg/g, and *Ziziphus nummularia* with 2.596 mg/g. In contrast, the plants with the least carotenoid content in this zone were *Koelaria macarantha* at 1.939 mg/g, *Cynodon dactylon* at 1.961 mg/g, and *Chenopodium album* at 1.949 mg/g. In the Less Polluted Zone (LPZ), which served as the control area, the leading carotenoid levels were found in *Arundo donax* with a concentration of 3.570 mg/g. Other notable contents were observed in *Cynodon dactylon* at 2.072 mg/g, *Coronops didymus* at 2.063 mg/g, *Persiaria glabra* at 2.011 mg/g, *Koelaria macarantha* at 1.851 mg/g, and *Achyranthes aspera* at 1.936 mg/g. The lower spectrum in the LPZ comprised *Dysphania ambrosioides* at just 0.511 mg/g, *Ziziphus nummularia* at 1.274 mg/g, *Cannabis sativa* at 1.281 mg/g, *Malvastrum coromandelianum* at 1.437 mg/g, and *Malava neglecta* at 1.604 mg/g. These detailed carotenoid levels within the respective pollution zones are documented in Figure 5.3 and Table 5.5.



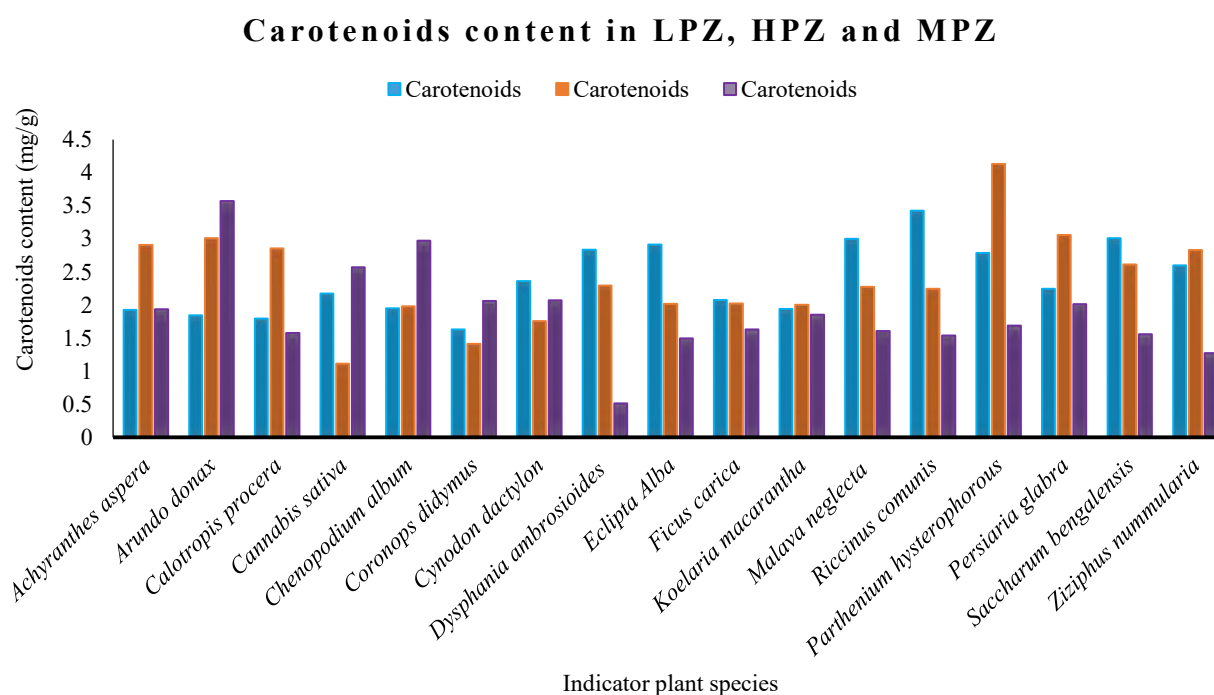


Figure 5. 3 Carotenoid content (mg/g) of the identified indicators plant species present in LPZ, MPZ, and HPZ of the IPWE.

### 5.3.8 Proline

The proline content in indicator plants has been observed to correlate directly with heavy metal pollution levels. The most significant proline accumulation was detected in the High Polluted Zone (HPZ) vegetation, with markedly higher levels than those found in the control environment or the Less Polluted Zone (LPZ), as illustrated in Figure 5.4 and Table 5.5. In the HPZ, *Coronops didymus* exhibited the highest proline concentration, reaching 1.994 mg/g, followed by *Cynodon dactylon* at 1.934 mg/g, *Chenopodium album* at 1.640 mg/g, and *Malvastrum coromandelianum* at 1.3453 mg/g. On the lower end of the spectrum within the same zone, *Koelaria macarantha* and *Ficus carica* showed the smallest proline accumulations, with levels at 1.027 mg/g and 1.064 mg/g, respectively, and *Saccharum bengalensis* at 1.239 mg/g. For plants in the moderately polluted zones, the maximum proline contents were found in *Chenopodium album* (1.969 mg/g), *Coronops didymus* (1.968 mg/g), and *Parthenium hysterophorus* (1.936 mg/g). *Malvastrum coromandelianum* and *Achyranthes aspera* also showed significant proline levels at 1.602 mg/g and 1.524 mg/g, respectively. In contrast, the minimum proline accumulation was observed in *Koelaria macarantha* at 0.408 mg/g and *Ficus carica* at 0.854 mg/g. Furthermore, in the LPZ, which experiences less pollution, *Malvastrum coromandelianum* had the highest recorded proline content at 1.440 mg/g, with *Arundo donax* and *Achyranthes aspera* also showing elevated



levels at 1.221 mg/g and 1.2734 mg/g, respectively. In terms of the lowest concentrations, *Koelaria macarantha*, *Ziziphus nummularia*, and *Calotropis procera* were on the minimal end, with proline contents of 0.234 mg/g, 0.620 mg/g, and 0.7942 mg/g accordingly. These findings in LPZ further underscore the trend that lower pollution correlates with reduced proline accumulation in plants.

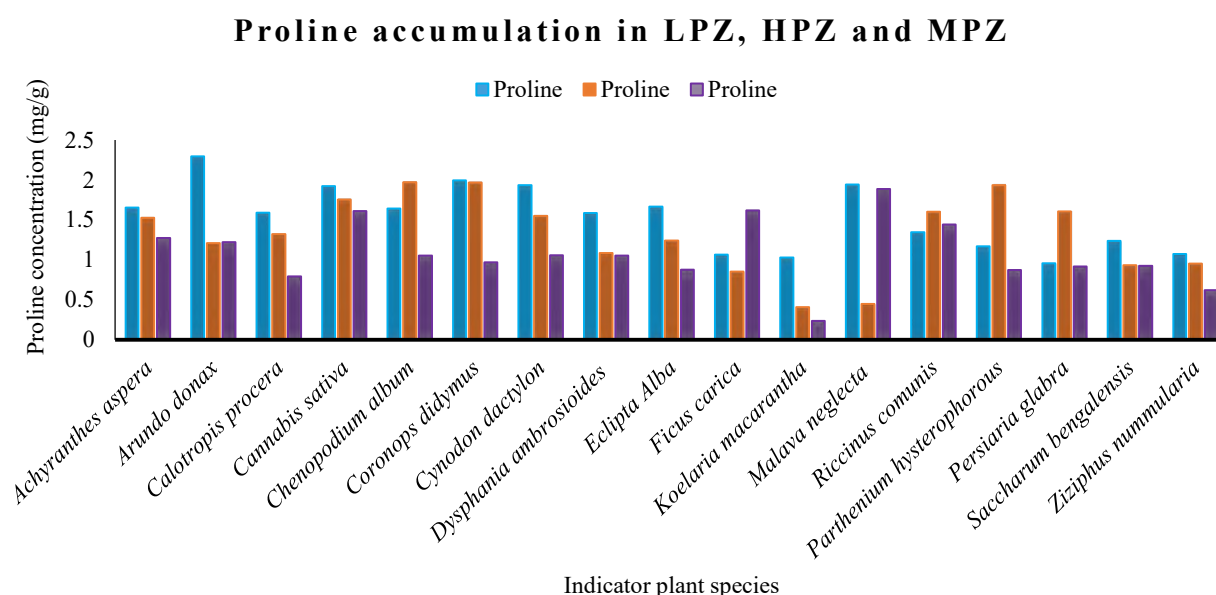


Figure 5.4 Proline accumulation (mg/g) of the identified indicators plant species present in LPZ, MPZ and HPZ of the IPWE.

Table 5.5 Detailed description of Chlorophyll-a, b and total carotenoids determined in the indicator plant species of three distinct zones.

Plants Names	Chlorophyll-a			Chlorophyll-b			Carotenoids			Proline		
	HPZ	MPZ	LPZ	HPZ	MPZ	LPZ	HPZ	MPZ	LPZ	HPZ	MPZ	LPZ
<i>Achyranthes aspera</i>	0.649	0.724	0.743	0.227	0.398	0.574	1.928	2.910	1.936	1.656	1.524	1.2734
<i>Arundo donax</i>	0.460	0.750	1.598	0.181	0.352	1.3614	1.841	3.008	3.570	2.295	1.2105	1.221
<i>Calotropis procera</i>	0.449	0.718	0.620	0.171	0.364	1.465	1.797	2.853	1.576	1.591	1.320	0.7942
<i>Cannabis sativa</i>	0.609	0.705	1.501	0.175	0.343	0.450	2.171	1.114	2.568	1.922	1.754	1.6114
<i>Chenopodium album</i>	0.419	0.580	1.102	0.2650	0.312	0.986	1.949	1.977	2.972	1.640	1.969	1.053
<i>Coronops didymus</i>	0.410	0.580	0.662	0.167	0.279	0.840	1.632	1.414	2.063	1.994	1.968	0.969
<i>Cynodon dactylon</i>	0.586	0.554	0.767	0.301	0.262	0.349	2.360	1.755	2.072	1.934	1.549	1.0568
<i>Dysphania ambrosioides</i>	0.709	0.520	0.916	0.292	0.281	0.450	2.836	2.293	0.511	1.584	1.084	1.0542



<i>Eclipta Alba</i>	0.721	0.502	0.760	0.3401	0.271	0.735	2.912	2.019	1.493	1.668	1.243	0.875
<i>Ficus carica</i>	0.677	0.527	1.508	0.512	0.298	0.648	2.074	2.024	1.632	1.064	0.854	1.6162
<i>Koelaria macarantha</i>	0.678	0.485	2.586	0.4721	0.346	0.469	1.939	2.004	1.851	1.027	0.408	0.234
<i>Malava neglecta</i>	0.543	0.655	1.584	0.4067	0.509	1.188	2.999	2.274	1.604	1.943	0.448	1.8847
<i>Riccinus comunis</i>	0.508	0.989	1.548	0.384	0.843	1.8842	3.427	2.244	1.537	1.3453	1.602	1.440
<i>Parthenium hysterophorous</i>	0.559	1.336	1.567	0.234	1.189	2.166	2.787	4.13	1.687	1.1703	1.936	0.8708
<i>Persiaria glabra</i>	0.569	1.566	2.557	0.266	1.4149	2.227	2.243	3.057	2.011	0.955	1.606	0.9171
<i>Saccharum bengalensis</i>	0.560	1.441	1.640	0.2723	1.294	1.2123	3.007	2.612	1.556	1.239	0.933	0.9260
<i>Ziziphus nummularia</i>	0.682	1.005	1.648	0.342	0.8587	0.8026	2.596	2.832	1.274	1.071	0.953	0.620

Table 5.6 Summary statistics of Bioaccumulation Coefficient (BAC), Translocation Factor (TF), Biological Concentration Factor (BCF), proline, chlorophyll-a, chlorophyll-b and carotenoids.

	BCF	TF	BAC	Proline	Chlorophyll-a	Chlorophyll-b	Carotenoids
<b>Chromium (Cr)</b>							
<b>LPZ</b>							
<b>Means</b>	1.0599	0.935	0.9923	1.862	0.484	0.2073	1.9374
<b>SD</b>	0.1016	0.074	0.1325	0.60	0.282	0.1743	1.130
<b>Maximum</b>	0.842	0.837	0.784	0.917	0.127	0.780	5.791
<b>Minimum</b>	0.8420	0.837	0.7840	0.3785	0.917	0.127	0.510
<b>HPZ</b>							
<b>Means</b>	1.034	0.884	0.9139	1.571	0.649	0.3575	2.596
<b>SD</b>	0.096	0.056	0.088	0.7071	0.355	0.279	1.421
<b>Maximum</b>	1.182	0.942	1.098	2.785	1.776	1.060	7.107
<b>Minimum</b>	0.929	0.735	0.822	0.9081	0.252	0.0670	1.011
<b>HPZ</b>							
<b>Means</b>	0.3763	0.652	0.277	0.881	0.738	0.558	2.263
<b>SD</b>	0.397	0.122	0.341	0.594	0.387	0.410	0.391
<b>Maximum</b>	1.179	0.927	1.028	2.763	2.763	1.424	3.068
<b>Minimum</b>	0.0719	0.4764	0.051	0.378	0.441	0.199	1.813
<b>Nickel (Ni)</b>							
<b>LPZ</b>							
<b>Means</b>	0.668	0.914	0.5798	1.258	0.9132	0.4425	2.334
<b>SD</b>	0.3921	0.8910	0.3884	0.6234	0.6071	0.4275	1.1032
<b>Maximum</b>	1.438	6.373	1.1680	2.971	3.55	1.778	7.1074
<b>Minimum</b>	0.0120	0.0333	0.00073	0.5092	0.2527	0.0626	0.5109



<b>HPZ</b>							
<b>Means</b>	0.640	1.23	0.606	1.581	1.312	0.470	2.225
<b>SD</b>	0.437	1.457	0.430	0.668	0.739	0.525	1.177
<b>Maximum</b>	1.012	6.373	0.99	2.97	3.55	1.778	5.79
<b>Minimum</b>	0.0120	0.033	0.000	0.645	0.444	0.062	0.510
<b>MPZ</b>							
<b>Means</b>	0.497	0.673	0.384	0.923	0.839	0.569	2.661
<b>SD</b>	0.437	0.388	0.396	0.497	0.351	0.410	0.5775
<b>Maximum</b>	1.438	1.198	1.1680	2.529	1.57	1.424	4.128
<b>Minimum</b>	0.015	0.0461	0.001	0.509	0.441	0.142	1.813
<b>Copper (Cu)</b>							
<b>LPZ</b>							
<b>Means</b>	0.971	0.765	0.8191	1.53	0.649	0.402	2.16
<b>SD</b>	0.309	0.340	0.442	0.690	0.499	0.482	1.181
<b>Maximum</b>	1.4465	0.978	1.29	1.924	1.924	1.778	5.791
<b>Minimum</b>	0.423	0.0142	0.046	0.0387	0.5042	0.062	0.510
<b>HPZ</b>							
<b>Means</b>	1.094	0.902	0.993	1.562	0.5651	0.309	2.26
<b>SD</b>	0.218	0.090	0.240	0.866	0.353	0.292	1.415
<b>Maximum</b>	1.446	1.048	1.304	2.785	1.77	1.060	7.107
<b>Minimum</b>	0.771	0.723	0.647	-0.9081	0.2527	0.064	1.011
<b>MPZ</b>							
<b>Means</b>	0.6002	0.783	0.493	1.24	0.803	0.558	2.51
<b>SD</b>	0.329	0.491	0.376	0.6799	0.368	0.415	0.565
<b>Maximum</b>	1.115	1.86	1.198	2.171	1.57	1.424	4.128
<b>Minimum</b>	0.003	0.0142	0.0005	0.102	0.441	0.1421	1.81
<b>Cadmium (Cd)</b>							
<b>LPZ</b>							
<b>Means</b>	1.333	1.679	1.997	1.862	0.751	0.525	2.230
<b>SD</b>	4.455	1.539	2.826	0.7491	0.558	0.5603	1.180
<b>Maximum</b>	26.001	10.333	18.629	2.785	1.924	1.778	7.1074
<b>Minimum</b>	0.1542	0.0133	0.0321	0.378	0.127	0.0626	0.510
<b>HPZ</b>							
<b>Means</b>	2.118	1.143	1.579	1.438	0.711	0.4621	2.353
<b>SD</b>	0.732	1.539	2.826	4.33	0.600	0.443	1.1049
<b>Maximum</b>	2.962	10.333	18.62	2.785	1.924	1.778	5.791
<b>Minimum</b>	0.1542	0.0133	0.0321	0.378	0.127	0.0626	0.510
<b>MPZ</b>							
<b>Means</b>	1.1218	0.143	0.579	0.438	1.711	0.8621	2.383
<b>SD</b>	0.7323	1.523	2.8456	4.234	0.6670	0.1234	1.123
<b>Maximum</b>	2.962	10.333	18.62	2.785	1.924	1.778	5.791
<b>Minimum</b>	0.1542	0.0133	0.0321	0.378	0.127	0.0626	0.510
<b>Zinc (Zn)</b>							
<b>LPZ</b>							
<b>Means</b>	0.82	0.88	0.76	1.59	0.59	0.25	2.18
<b>SD</b>	0.593	0.332	0.437	0.795	0.712	0.562	1.237
<b>Maximum</b>	2.00	1.00	1.00	3.00	2.00	2.00	6.00
<b>Minimum</b>	0.001	0.002	0.003	1.004	0.005	0.002	1.005



HPZ							
Means	1.006	1.003	1.004	1.004	1.29	0.47	0.18
SD	0.000	0.000	0.000	0.47	0.624	0.393	1.33
Maximum	1.004	1.04	1.06	2.05	2.04	1.04	7.9
Minimum	1.005	1.87	1.05	1.034	0.0065	0.007	1.09
MPZ							
Means	0.508	0.706	0.438	0.929	0.803	0.558	2.51
SD	0.372	0.406	0.400	0.495	0.368	0.415	0.565
Maximum	0.99	1.40	0.96	2.52	1.57	1.42	4.12
Minimum	0.015	0.003	0.0014	0.509	0.441	0.141	1.88
Lead (Pb)							
LPZ							
Means	0.71	0.82	0.71	1.53	0.65	0.29	2.29
SD	0.470	0.393	0.470	0.514	0.702	0.588	1.121
Maximum	1.098	1.08	1.65	2.96	2.96	2.06	6.07
Minimum	0.07	0.097	0.07	1.97	0.097	0.097	1.097
HPZ							
Means	0.88	1.097	0.88	1.53	0.41	0.12	2.12
SD	0.332	0.000	0.332	0.874	0.618	0.332	1.409
Maximum	1.0064	1.097	1.086	3.0765	2.086	1.086	7.094
Minimum	0.0765	1.0765	0.0986	-1.085	0.0654	0.0543	1.954
MPZ							
Means	0.65	0.65	0.47	1.29	1.00	0.35	2.59
SD	0.493	0.493	0.514	0.774	0.500	0.493	0.618
Maximum	1.0654	1.987	1.876	2.654	2.876	1.543	4.654
Minimum	0.986	0.909	0.543	0.064	0.075	0.085	2.054
Arsenic (As)							
LPZ							
Means	0.035	0.024	0.057	0.180	0.002	0.057	1.099
SD	0.093	0.083	0.0314	0.034	0.004	0.900	0.008
Maximum	0.3654	0.087	0.876	0.654	0.073	0.040	1.604
Minimum	0.036	0.059	0.003	0.062	0.003	0.035	0.076
HPZ							
Means	0.45	0.24	0.057	0.180	0.02	0.57	1.099
SD	0.93	0.83	0.314	0.034	0.04	0.900	0.08
Maximum	1.3654	1.087	1.876	1.654	1.073	1.040	1.604
Minimum	0.336	0.259	0.033	0.262	0.043	0.135	0.176
MPZ							
Means	0.121	0.212	0.121	0.53	0.325	0.139	1.139
SD	0.8570	0.3493	0.5670	0.3414	0.2102	0.678	0.3221
Maximum	1.098	1.08	1.65	2.96	2.96	2.06	6.07
Minimum	0.07	0.097	0.07	1.97	0.097	0.097	1.097
Mercury (Hg)							
LPZ							
Means	0.333	0.679	0.347	0.562	0.751	0.525	0.230
SD	0.455	0.539	0.826	0.7491	0.558	0.5603	1.180
Maximum	1.001	1.333	1.629	1.785	0.924	0.778	1.1074
Minimum	0.1542	0.0133	0.0321	0.378	0.127	0.0626	0.510



HPZ							
<b>Means</b>	0.436	0.703	0.904	0.904	0.874	0.675	0.321
<b>SD</b>	0.0342	0.453	0.0324	0.4342	0.62453	0.3943	1.3323
<b>Maximum</b>	1.032	1.0423	1.0641	2.0543	2.0443	1.0454	2.921
<b>Minimum</b>	0.975	0.8732	0.432	0.5421	0.6313	0.43122	0.9712
MPZ							
<b>Means</b>	0.26	0.160	0.240	0.2129	0.213	0.212	0.259
<b>SD</b>	0.0930	0.0433	0.0324	0.0740	0.4301	0.09312	0.0181
<b>Maximum</b>	1.654	1.870	1.8753	0.6540	0.8761	0.54321	0.65432
<b>Minimum</b>	0.936	0.2319	0.3443	0.3164	0.1275	0.1185	0.9954

### 5.3.9 Phytoremediation Efficacy in Indicator Plants: Insights from Probit and OLS

#### Regression Models

The Probit and Ordinary Least Squares (OLS) regression models were employed to investigate the phytoremediation capacity of indicator plants in response to various heavy metal contaminants. The Probit model established the significance of proline as a reliable biomarker for phytoremediation success. This was particularly evident with metals like Chromium (Cr) and Lead (Pb), where proline coefficients were notably high and positive ( $\beta$  0.12 &  $p = 0.0236$  and  $\beta$  0.864 &  $p\text{-value} = 0.8754$ , respectively), indicating a strong likelihood of effective phytoremediation. In contrast, chlorophyll-a levels showed a negative correlation across all metals, suggesting an inverse relationship with phytoremediation efficacy, as seen with coefficients like ( $\beta$  -0.231 &  $p\text{-value} = 0.0125$ ) for Cr and ( $\beta$  -0.826 &  $p\text{-value} = 0.0422$ ) for Mercury (Hg). Other physiological markers, such as chlorophyll-b and carotenoids, displayed varied responses, indicating their complex roles in plant adaptation to heavy metal exposure (Table 5.7).

In parallel, the OLS regression model focused on the specific impacts of heavy metals on plant physiology. It confirmed the positive correlation between proline levels and effective phytoremediation, especially for chromium, and highlighted the negative correlation of chlorophyll-a and chlorophyll-b levels with heavy metal stress. For example, in chromium exposure, proline levels had a positive and significant impact ( $\beta$  0.36 &  $p\text{-value} = 0.000$ ), while chlorophyll-a and chlorophyll-b exhibited a negative and significant relationship ( $\beta$  -0.0123 &  $p\text{-value} = 0.053$  and  $\beta$  -0.0456 &  $p\text{-value} = 0.033$ , respectively). Similarly, for other metals like nickel, copper, cadmium, zinc, and lead, proline levels were consistently positively associated with phytoremediation efficiency, whereas chlorophyll-a and chlorophyll-b mostly presented negative correlations. Carotenoids, however, did not display a



significant correlation in most cases, except for a positive relationship with zinc phytoremediation (Table 5.7 and Figure 5.5 (a-h)).

Overall, these models offer a comprehensive understanding of the key physiological changes in plants, particularly highlighting the roles of proline and chlorophyll concentrations as crucial indicators of phytoremediation potential under heavy metal stress. The coherence of these findings across various metals and physiological markers validates the study and deepens the understanding of plant adaptive responses in polluted environments.

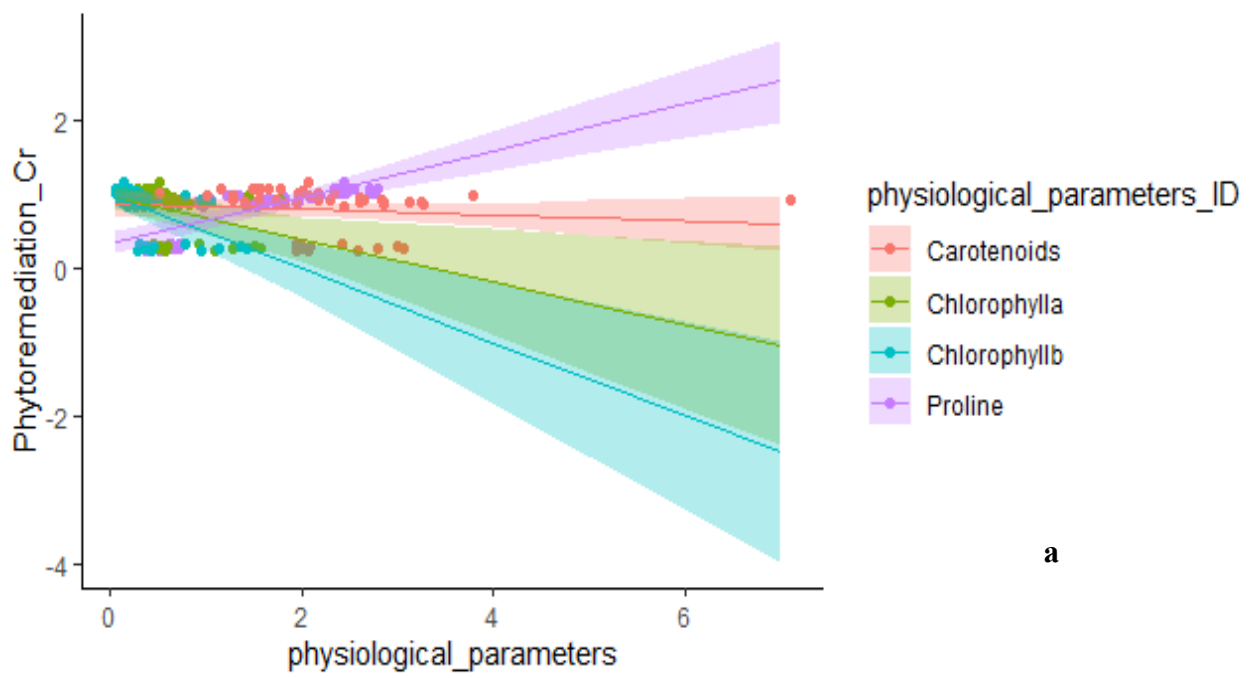
Table 5.7 Ordinary least square (OLS) and Probit Model.

<b>Dependent Variable is Phytoremediation (Cr)</b>						
<b>Variables</b>	<b>OLS Model</b>			<b>Probit model</b>		
	<b>Coefficient</b>	<b>S.E</b>	<b>p-value</b>	<b>Coefficient</b>	<b>S.E</b>	<b>p-value</b>
Proline	<b>0.36***</b>	0.0018	0.000	<b>0.12**</b>	0.03012	0.0236
Chlorophyll-a	<b>-0.0123*</b>	0.2145	0.053	<b>-0.231**</b>	0.0134	0.0125
Chlorophyll-b	<b>-0.0456**</b>	0.0023	0.033	<b>-0.089*</b>	0.4943	0.0632
Carotenoids	0.034	0.3421	0.341	<b>0.325*</b>	0.3341	0.0735
<b>Dependent Variable is Phytoremediation (Ni)</b>						
Proline	<b>0.016*</b>	0.0231	0.0589	0.124	0.0037	0.4372
Chlorophyll-a	<b>-0.11*</b>	0.0055	0.0753	<b>-0.451**</b>	0.0035	0.0245
Chlorophyll-b	<b>-0.06**</b>	0.0013	0.0245	<b>-0.081*</b>	0.1143	0.0678
Carotenoids	<b>0.114**</b>	0.0021	0.0437	<b>0.325*</b>	0.0301	0.0864
<b>Dependent Variable is Phytoremediation (Cu)</b>						
Proline	<b>0.22**</b>	0.0016	0.0476	<b>0.11***</b>	0.7013	0.000
Chlorophyll-a	<b>-0.03***</b>	0.0045	0.000	<b>-0.201**</b>	0.7132	0.0265
Chlorophyll-b	-0.0126	0.0123	0.8734	<b>-0.035**</b>	0.4740	0.0496
Carotenoids	<b>0.0124*</b>	0.0021	0.0759	<b>0.021*</b>	0.6041	0.7435
<b>Dependent Variable is Phytoremediation (Cd)</b>						
Proline	0.45	0.0103	0.9462	<b>0.0892*</b>	0.01026	0.0637
Chlorophyll-a	<b>-0.0134**</b>	0.2045	0.03659	<b>-0.03***</b>	0.0170	0.000
Chlorophyll-b	-0.0345	0.1103	0.8462	<b>-0.008**</b>	0.1043	0.0368
Carotenoids	<b>0.011**</b>	0.3031	0.0683	<b>0.12*</b>	0.2511	0.0649
<b>Dependent Variable is Phytoremediation (Zn)</b>						

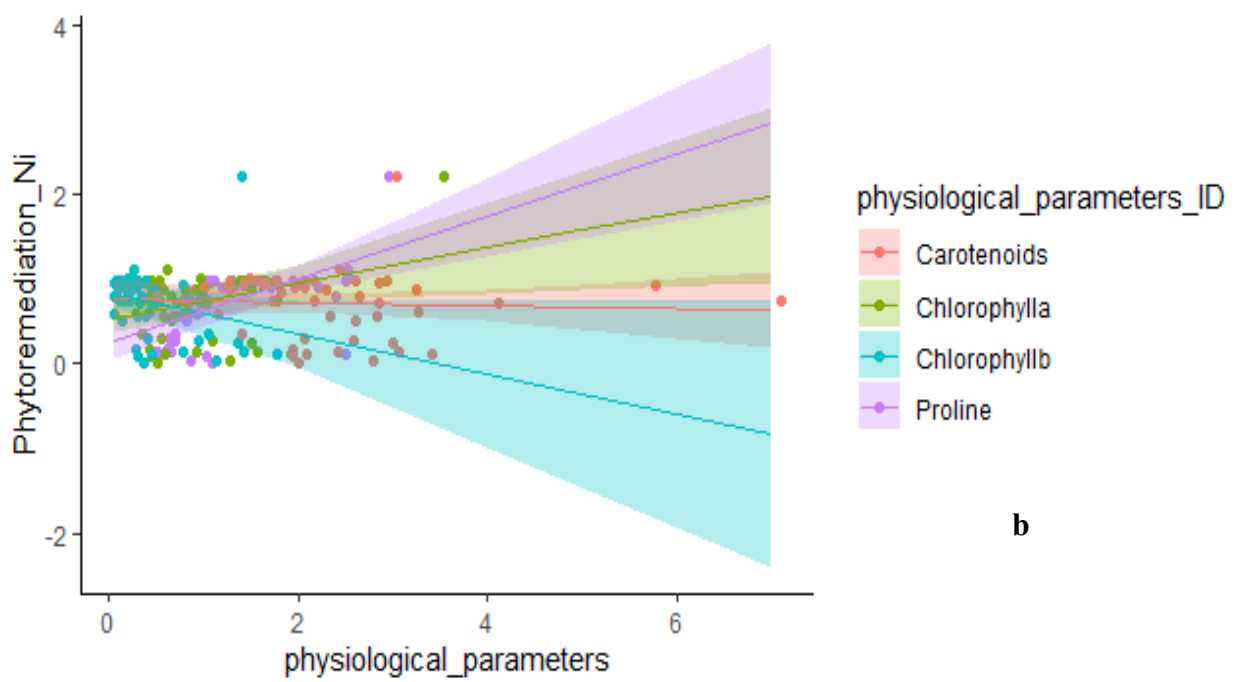


Proline	<b>0.012*</b>	0.0023	0.0854	<b>0.0345**</b>	0.0011	0.0437
Chlorophyll-a	<b>-0.0100***</b>	0.2145	0.000	<b>-0.02***</b>	0.0454	0.000
Chlorophyll-b	<b>-0.560***</b>	0.0986	0.000	<b>-0.007*</b>	0.4708	0.0762
Carotenoids	<b>0.034*</b>	0.3020	0.0749	<b>0.0025*</b>	0.0943	0.05639
<b>Dependent Variable is Phytoremediation (Pb)</b>						
Proline	<b>0.059**</b>	0.076	0.0356	0.864	0.00543	0.8754
Chlorophyll-a	<b>-0.097**</b>	0.2145	0.0427	<b>-0.031**</b>	0.0194	0.0433
Chlorophyll-b	<b>-0.0087**</b>	0.074	0.0368	<b>-0.0214*</b>	0.0631	0.0789
Carotenoids	<b>0.021**</b>	0.3703	0.0447	<b>0.3021*</b>	0.0045	0.0655
<b>Dependent Variable is Phytoremediation (As)</b>						
Proline	<b>0.123**</b>	0.0143	0.0116	<b>0.0824*</b>	0.0642	0.0548
Chlorophyll-a	-0.102	0.0713	0.7463	-0.623	0.1254	0.7534
Chlorophyll-b	<b>-0.208**</b>	0.634	0.0331	-0.178	0.7653	0.9147
Carotenoids	0.291	0.0934	0.8463	0.1003	0.0543	0.4765
<b>Dependent Variable is Phytoremediation (Hg)</b>						
Proline	<b>0.097**</b>	0.6513	0.0226	<b>0.0824**</b>	0.00543	0.0344
Chlorophyll-a	<b>-0.164**</b>	0.0928	0.0437	<b>-0.826**</b>	0.0019	0.0422
Chlorophyll-b	<b>-0.1092**</b>	0.02714	0.0346	<b>-0.2017*</b>	0.00982	0.0586
Carotenoids	<b>0.0265**</b>	0.3421	0.0438	<b>0.320*</b>	0.0915	0.0684



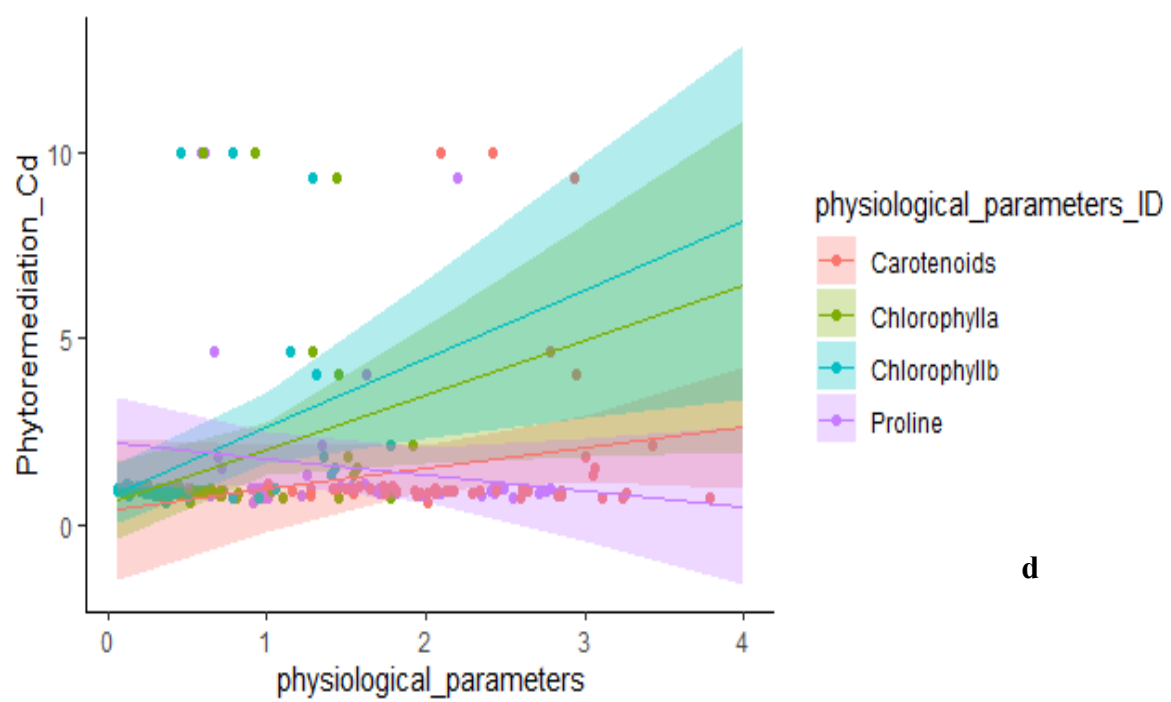
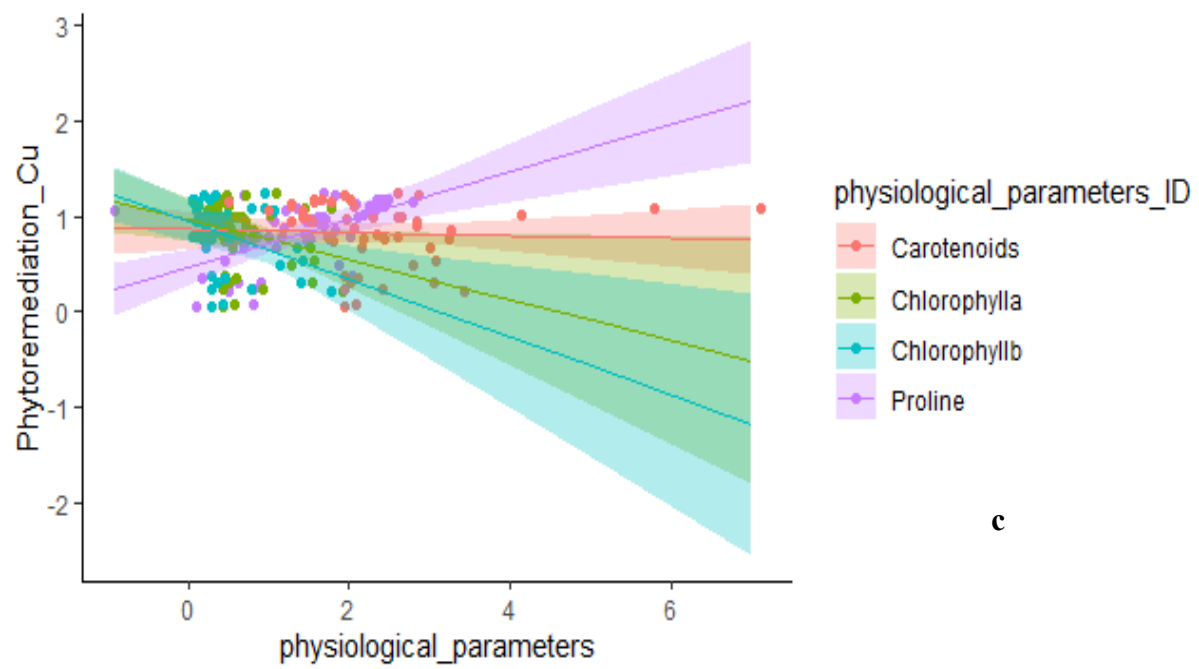


**a**

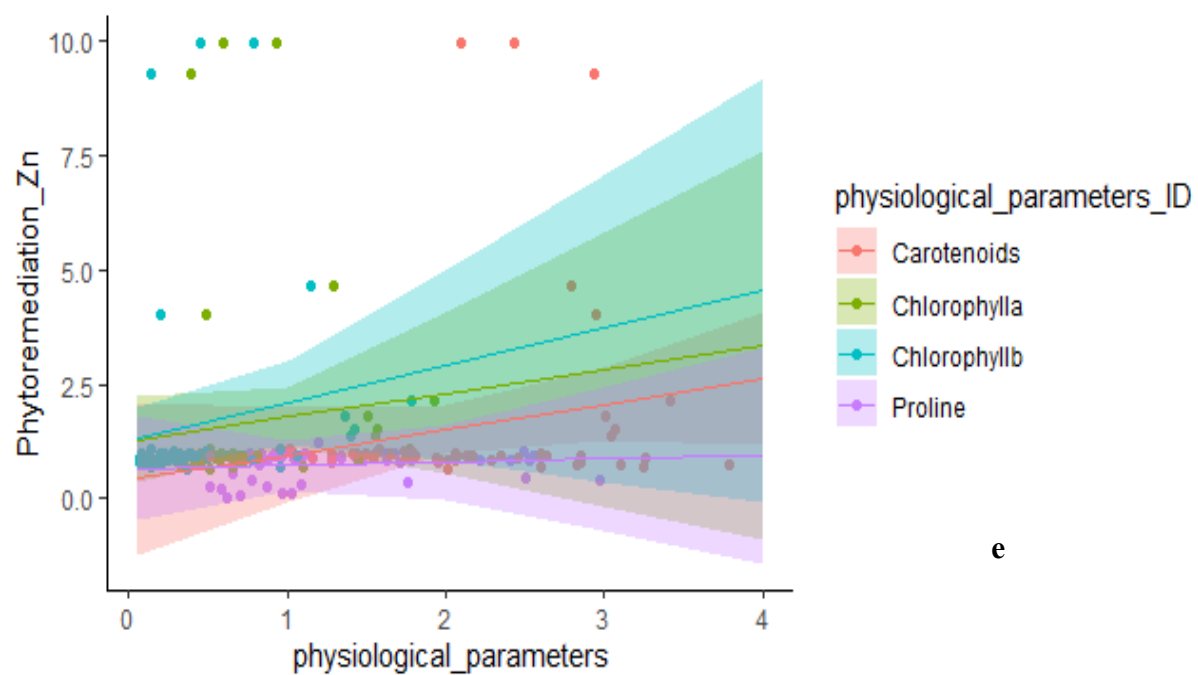


**b**

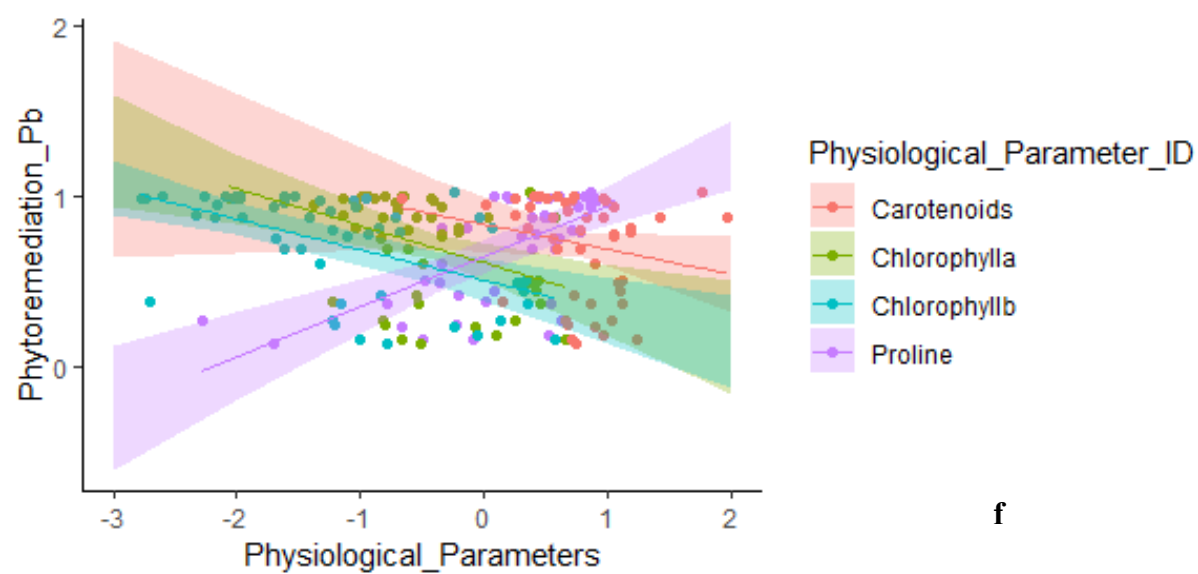








**e**



**f**



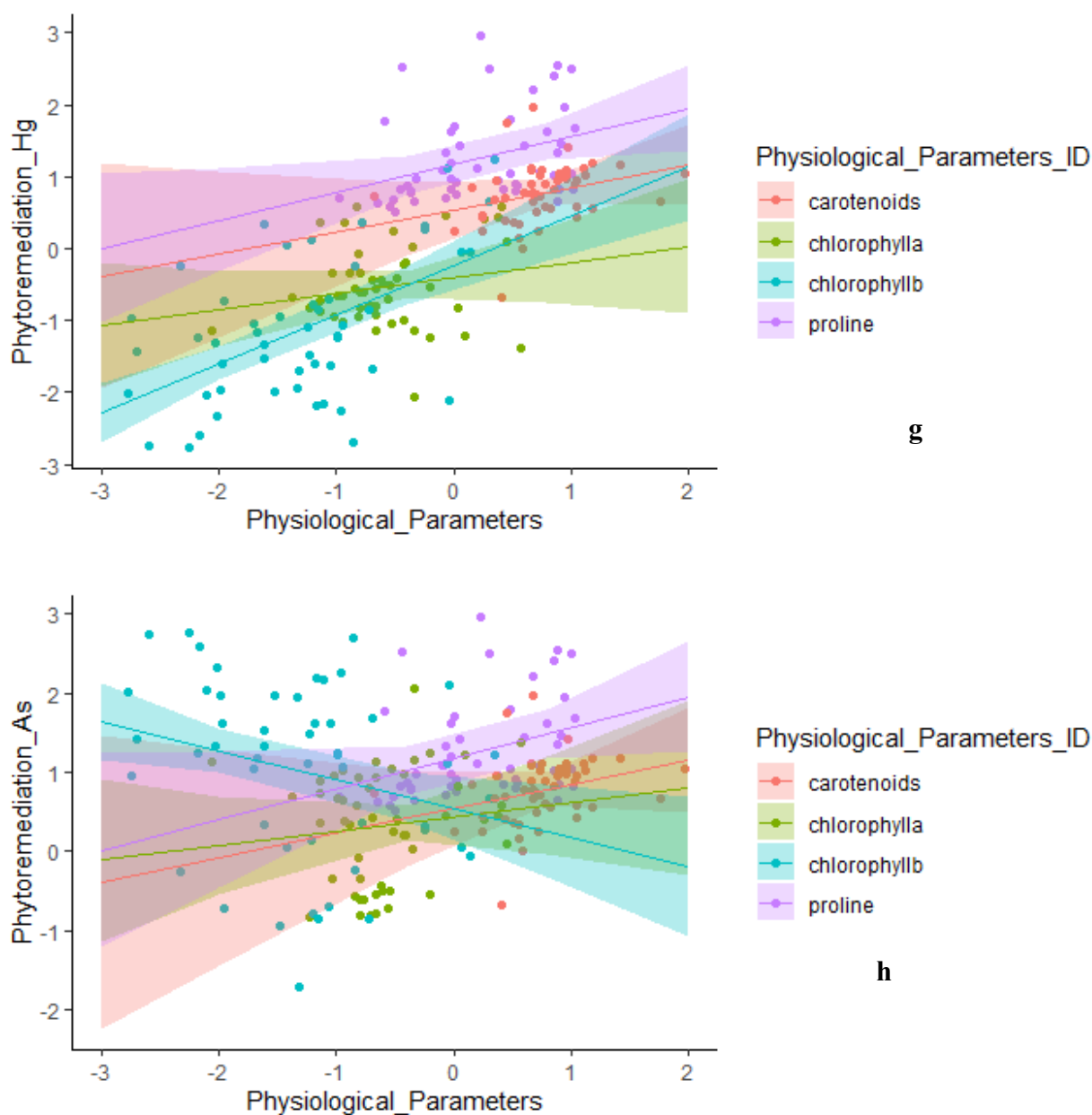


Figure 5.5 (a-h) Physiological Impact on Indicator Plants from Heavy Metal Accumulation (Cr, Ni, Cu, Cd, Zn, Pb, Hg, As).

### 5.3.10 Impact of Heavy Metal Remediation on Plant Physiology across Different Polluted Zones: A Generalized Mixed-Effect Modeling Approach

Generalized Mixed Effects Modeling (GLMM) assessed how individual heavy metals affect physiological changes or adaptive responses in the indicator plant species within the industrially polluted zone. This modeling approach simultaneously analyzed fixed and random effects. Our model treated proline, chlorophyll-a, and chlorophyll-b as fixed effects, whereas phyto remediation represented a random effect. The analysis revealed a notably positive correlation between the phyto remediation of chromium and the accumulation of



proline (0.17655) in the indicator species. Conversely, chlorophyll-a and chlorophyll-b showed a significantly negative correlation with chromium phytoremediation, with coefficients of -0.28349 and -0.35413, respectively (Table 5.8 and fig. 5.6). The interspecific variance recorded was 21% between the low pollution zone (LPZ) and high pollution zone (HPZ) and 2% between the medium pollution zone (MPZ) and LPZ. The total observed variation among chromium and other heavy metals was 24% (Table 5.8 and fig. 5.6). The analysis using Generalized Mixed Effects Modeling revealed that nickel phytoremediation by the selected indicator species was significantly and positively correlated with the accumulation of proline, with a correlation coefficient of 0.13525.

In contrast, chlorophyll-a and chlorophyll-b levels demonstrated significant negative correlations with nickel phytoremediation, with coefficients of -0.73382 and -0.96901, respectively. The analysis also highlighted interspecific variance, with an 11% difference observed between the high pollution zone (HPZ) and the low pollution zone (LPZ) and a 29% variance noted between nickel and other metals (Table 3.7). The analysis indicated an insignificant positive correlation (0.117) between copper phytoremediation and proline accumulation in the indicator plants. Chlorophyll-a levels were insignificantly correlated (0.10478) with copper phytoremediation, whereas chlorophyll-b levels had a significant negative correlation (-0.37389) with the phytoremediation of copper. Interspecific variance was noted at different percentages across pollution zones: 7% between the high pollution zone (HPZ) and other zones, 9% between the medium pollution zone (MPZ) and others, and 1% between the low pollution zone (LPZ) and others. The overall variance between copper and the other metals analyzed was 13% (Table 5.8). The relationship between cadmium removal by the indicators and proline increase was strongly positive, with a correlation value of 2.1747. On the contrary, the analysis indicated a robust negative impact on chlorophyll-a and chlorophyll-b levels due to cadmium removal activities, with values of -4.1554 and -0.6343, respectively. Disparities in species reactions across various contamination levels were recorded at 5%, 8%, and 17% for low, medium, and high pollution areas. The overall difference in response to cadmium compared to other metals was 28% (Table 5.8).

Proline accumulation in the indicator plants within the polluted zone has been found to have a positive correlation (0.18806) with the phytoremediation of zinc. Initially, chlorophyll-a levels appeared to rise with an increase in zinc remediation within a controlled environment; however, this trend reverses, showing a negative relationship in areas of high pollution. Chlorophyll-b concentration, in contrast, consistently demonstrates a significant



negative correlation (-0.4506) with zinc phytoremediation. Variability between the highly polluted areas and other contaminated regions was 5%, 2%, and 9%, respectively. When comparing the interaction between zinc and other potentially toxic elements, the variance was established at 18% (as detailed in Table 5.8). There is a significant rise in proline accumulation (0.50535) in correlation with the increase in lead phytoremediation efforts. Concurrently, chlorophyll-a and chlorophyll-b experience decreases (-0.40761 & -0.22875) as the accumulation of lead (Pb) in the plant species under study intensifies. Differences between the species in various pollution levels—polluted, medium pollution zones (MPZ), and low pollution zones (LPZ)—were noted as 4%, 6%, and 5%, respectively. The variation between the response to copper and other analyzed metals within these zones was identified to be 12% (Table 3.7). The model indicated that in arsenic detoxification by IPWE's indicator species, proline levels increased significantly (0.8200), whereas chlorophyll-a and chlorophyll-b exhibited significant decreases, with values of -0.034 and -0.045, respectively. Species variance across different degrees of polluted zones was considerable, with values of 10%, 13%, and 15%, and the variance for arsenic compared to other potentially toxic elements was measured at 25% (Table 5.8 and fig. 5.6).

In the case of mercury, the proline content in the indicator species showed a significant uptick (0.09453) due to phytoremediation activities, while chlorophyll-b levels decreased substantially (-0.576). Chlorophyll-a demonstrated an insignificant but positive trend (0.018) in the presence of mercury. Variability among species in low pollution zone (LPZ), medium pollution zone (MPZ), and high pollution zone (HPZ) was recorded at 4%, 3%, and 10%, respectively. The disparity between the reaction to magnesium and other metals was established at 14% (Table 5.8 and fig. 5.6 (A to H)).



Table 5.8 Generalized Mixed Effect Model

Heavy metal Cu																						
$Y_{BCF}$	$\alpha$	$\beta_{IVI}$	$\beta_{1\text{Carotenoids}}$	$\beta_{2\text{proline}}$	$\beta_{3\text{Chlorophyll-a}}$	$\beta_{4\text{Chlorophyll-b}}$	$b_{zone} + \mu_{ij}$	$Y_{TF}$	$\alpha$	$\beta_{IVI}$	$\beta_{1\text{Carotenoids}}$	$\beta_{2\text{proline}}$	$\beta_{3\text{Chlorophyll-a}}$	$\beta_4$	$Y_{BAC}$	$\alpha$	$\beta_{IVI}$	$\beta_{1\text{Carotenoids}}$	$\beta_{2\text{proline}}$	$\beta_{3\text{Chlorophyll-a}}$		



Chlorophyll-a	0.10478	0.38025	0.276	0.7842	0.47884	0.50215	0.954	0.3454	0.08612	0.48459	0.178	0.859
Chlorophyll-b	-0.37389	0.25394	-1.472	0.1482	-0.30428	0.33668	-0.904	0.3709	-0.43677	0.32406	-1.348	0.184
<b>Random effect</b>	<b>Variance</b>	<b>St. dev</b>			<b>Variance</b>	<b>St. dev</b>			<b>Variance</b>	<b>St. dev</b>		
All Zones	0.29752	0.772521			0.001735	0.03579			0.1342	0.3159		
Zone (LPZ-HPZ)	0.09136	0.031467			0.00022	0.000342			0.07146	0.056467		
Zone (HPZ-MPZ)	0.04164	0.042198			0.00011	0.000213			0.09316	0.077198		
Zone (MPZ-LPZ)	0.112523	0.143257			0.00231	0.05421			0.012576	0.113257		
Residual	0.06439	0.2538			0.1142	0.3379			0.10509	0.3242		

#### Heavy metal Cd

<b>Fixed effect</b>	<b>Coefficient</b>	<b>St. Error</b>	<b>t-value</b>	<b>p-value</b>	<b>Coefficient</b>	<b>St. Error</b>	<b>t-value</b>	<b>p-value</b>	<b>Coefficient</b>	<b>St. Error</b>	<b>t-value</b>	<b>p-value</b>
Intercept	3.0526	3.6765	0.83	0.4124	<b>2.1074*</b>	1.2135	1.737	0.0914	1.8796	2.0795	0.904	0.371
IVI	0.1011	0.6609	0.153	0.8791	-0.2948	0.2209	-1.334	0.1893	-0.1214	0.3768	-0.322	0.749
Carotenoids	0.9264	1.2904	0.718	0.4766	<b>-0.8401*</b>	0.4362	-1.926	0.0606	-1.0527	0.805	-1.308	0.198
Proline	<b>2.1747**</b>	0.8707	2.498	0.0163	0.3147	0.2941	1.07	0.0534	0.2466	0.5377	0.459	0.043
Chlorophyll-a	-9.136	6.5088	-1.404	0.1674	<b>4.1554*</b>	2.1978	1.891	0.0651	3.4736	4.026	0.863	0.393
Chlorophyll-b	4.4915	4.1889	1.072	0.2895	-0.6343	1.419	-0.447	0.6571	0.5399	2.6528	0.204	0.84
<b>Random effect</b>	<b>Variance</b>	<b>St. dev</b>			<b>Variance</b>	<b>St. dev</b>			<b>Variance</b>	<b>St. dev</b>		
All Zones	0.3916	1.979			0.2825	0.5315			0.0001	0.0003		
Zone (LPZ-HPZ)	0.073245	0.23146			0.056241	0.112755			0.00146	0.000467		
Zone (HPZ-MPZ)	0.094266	0.19356			0.085541	0.137431			0.00116	0.000198		
Zone (MPZ-LPZ)	0.218563	0.934158			0.17432	0.16732			0.002576	0.003257		
Residual	17.038	4.128			1.9587	1.3995			6.921	2.631		

#### Heavy metal Zn

<b>Fixed effect</b>	<b>Coefficient</b>	<b>St. Error</b>	<b>t-value</b>	<b>p-value</b>	<b>Coefficient</b>	<b>St. Error</b>	<b>t-value</b>	<b>p-value</b>	<b>Coefficient</b>	<b>St. Error</b>	<b>t-value</b>	<b>p-value</b>
Intercept	<b>0.6983***</b>	0.20788	3.359	0.00221	<b>0.45486*</b>	0.20076	2.266	0.03167	<b>0.50239*</b>	0.21116	2.379	0.023
IVI	0.01678	0.03496	0.48	0.63354	0.04534	0.03568	1.271	0.21461	0.02652	0.03628	0.731	0.468
Carotenoids	0.01445	0.0672	0.215	0.83073	0.08541	0.07375	1.158	0.25303	0.02326	0.07026	0.331	0.742
Proline	<b>0.18806***</b>	0.04338	4.335	0.00310	<b>0.16418**</b>	0.04725	3.475	0.0310	<b>0.20342***</b>	0.04532	4.488	0.000
Chlorophyll-a	0.27867	0.32614	0.854	0.39756	<b>0.7794*</b>	0.35927	2.169	0.03551	0.34452	0.34108	1.01	0.318
Chlorophyll-b	<b>-0.4506*</b>	0.21784	-2.068	0.04462	-0.3627	0.24072	-1.507	0.1391	<b>-0.45878*</b>	0.2279	-2.013	0.05
<b>Random effect</b>	<b>Variance</b>	<b>St. dev</b>			<b>Variance</b>	<b>St. dev</b>			<b>Variance</b>	<b>St. dev</b>		
All Zones	0.1832	0.9353			0.1153	0.3396			0.1347	0.4161		
Zone (LPZ-HPZ)	0.056572	0.25137			0.01627	0.02755			0.01462	0.04675		
Zone (HPZ-MPZ)	0.025428	0.153319			0.01556	0.007431			0.011644	0.019842		
Zone (MPZ-LPZ)	0.092419	0.257682			0.08435	0.11732			0.025761	0.325753		
Residual	0.04740	0.2177			0.058238	0.24132			0.05190	0.2278		

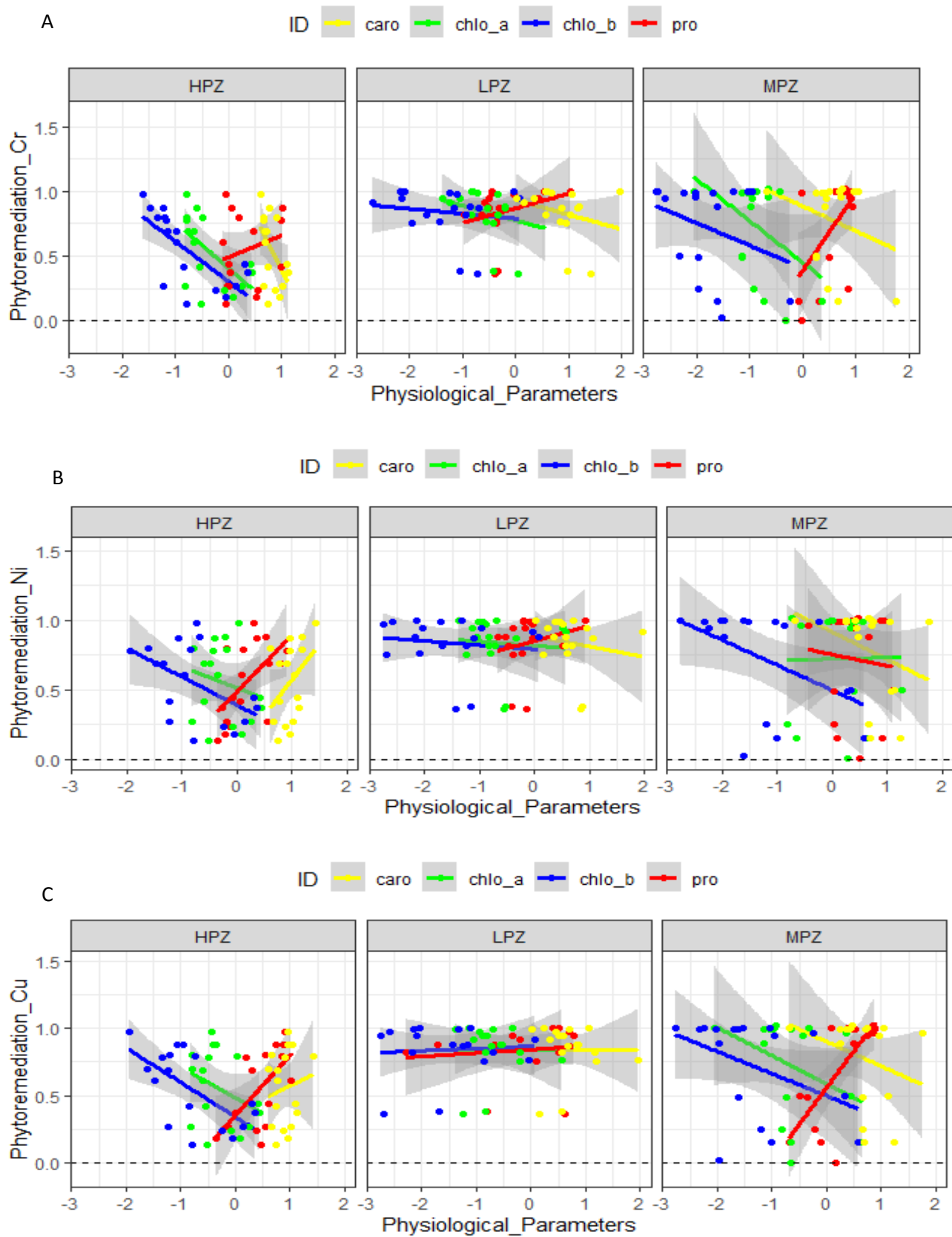
#### Heavy metal Pb

<b>Fixed effect</b>	<b>Coefficient</b>	<b>St. Error</b>	<b>t-value</b>	<b>p-value</b>	<b>Coefficient</b>	<b>St. Error</b>	<b>t-value</b>	<b>p-value</b>	<b>Coefficient</b>	<b>St. Error</b>	<b>t-value</b>	<b>p-value</b>
Intercept	<b>0.50535**</b>	0.1967	2.569	0.0146	<b>0.50639**</b>	0.21051	2.406	0.0221	<b>0.30293*</b>	0.24084	1.258	0.217

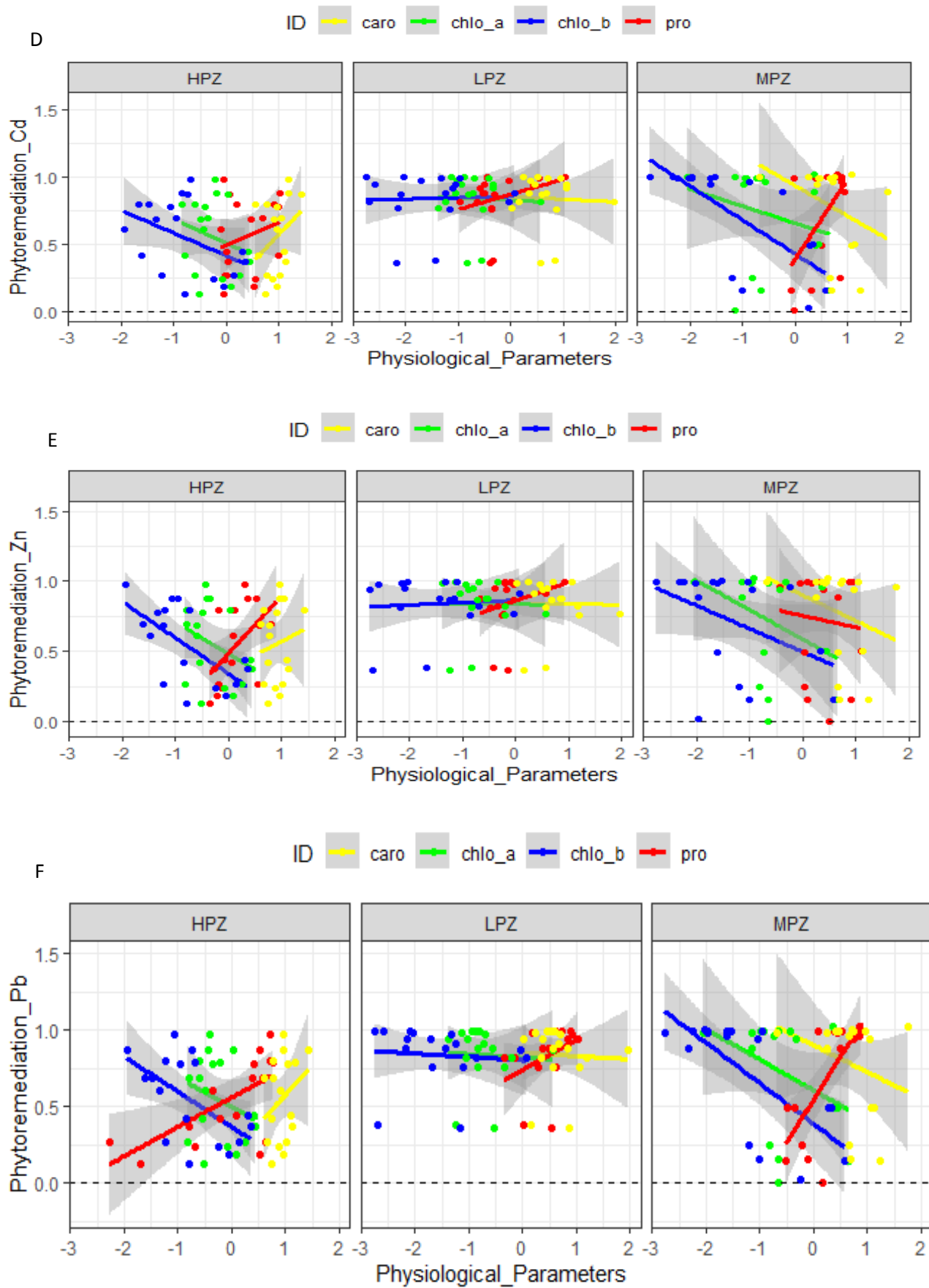


IVI	0.04671	0.03583	1.303	0.1995	0.05649	0.03868	1.46	0.1541	0.07139	0.04365	1.635	0.109
Carotenoids	<b>0.15797*</b>	0.06896	2.291	0.0268	-0.04379	0.07747	-0.565	0.5748	0.11661	0.08346	1.397	0.169
Proline	0.03053	0.04474	0.682	0.0045	<b>0.13275*</b>	0.05007	2.651	0.0023	<b>0.07305*</b>	0.05418	1.348	0.003
Chlorophyll-a	-0.40761	0.34896	-1.168	0.249	0.42567	0.39085	1.089	0.2819	-0.18809	0.42257	-0.445	0.658
Chlorophyll-b	-0.22875	0.23196	-0.986	0.3294	-0.30654	0.26037	-1.177	0.2453	-0.34521	0.28078	-1.229	0.225
<b>Random effect</b>	<b>Variance</b>	<b>St. dev</b>			<b>Variance</b>	<b>St. dev</b>			<b>Variance</b>	<b>St. dev</b>		
All Zones	0.06963	0.8345			0.02549	0.5049			0.012	0.1128		
Zone (LPZ-HPZ)	0.001436	0.00137			0.00159	0.02345			0.0044	0.0563		
Zone (HPZ-MPZ)	0.005411	0.003419			0.00556	0.007461			0.0063	0.0467		
Zone (MPZ-LPZ)	0.01040	0.036718			0.01445	0.07567			0.0057	0.0053		
Residual	0.049389	0.22224			0.062894	0.25079			0.0722	0.2688		
<b>Heavy metal As</b>												
<b>Fixed effect</b>	<b>Coefficient</b>	<b>St. Error</b>	<b>t-value</b>	<b>p-value</b>	<b>Coefficient</b>	<b>St. Error</b>	<b>t-value</b>	<b>p-value</b>	<b>Coefficient</b>	<b>St. Error</b>	<b>t-value</b>	<b>p-value</b>
Intercept	<b>0.8200**</b>	0.0234	3.815	0.0034	0.3421	0.0453	0.7563	0.5635	0.39563	0.56341	1.452	0.096
IVI	0.3452	0.0934	0.093	0.5634	0.07453	0.00834	0.9452	0.3214	0.86745	0.34526	1.345	0.856
Carotenoids	0.0194	0.7354	1.452	0.2413	0.1342	0.03542	-0.341	0.2135	0.4563	0.64524	1.356	0.085
Proline	<b>0.6341**</b>	0.75635	5.924	0.0031	<b>0.4659**</b>	0.06453	4.8341	0.0005	0.5645	0.64521	1.876	0.452
Chlorophyll-a	<b>-0.034*</b>	0.02461	-2.453	0.009	<b>-0.0342*</b>	0.08564	-3.561	0.041	-0.3524	0.56442	-0.325	0.785
Chlorophyll-b	<b>-0.045**</b>	0.01351	-3.563	0.0001	-0.4532	0.6231	-1.851	0.0845	-0.5644	0.43512	-1.351	0.354
<b>Random effect</b>	<b>Variance</b>	<b>St. dev</b>			<b>Variance</b>	<b>St. dev</b>			<b>Variance</b>	<b>St. dev</b>		
All Zones	0.25634	0.992347			0.2453	0.6745533			0.1124	0.4961541		
Zone (LPZ-HPZ)	0.10411	0.125892			0.01534	0.011487			0.04516	0.012249		
Zone (HPZ-MPZ)	0.134563	0.341371			0.055745	0.0024561			0.0063122	0.0993124		
Zone (MPZ-LPZ)	0.194582	0.369314			0.14667	0.2643219			0.0953217	0.124743		
Residual	0.4533	0.0923			0.00756	0.45331			0.07865	0.4634		
<b>Heavy metal Hg</b>												
<b>Fixed effect</b>	<b>Coefficient</b>	<b>St. Error</b>	<b>t-value</b>	<b>p-value</b>	<b>Coefficient</b>	<b>St. Error</b>	<b>t-value</b>	<b>p-value</b>	<b>Coefficient</b>	<b>St. Error</b>	<b>t-value</b>	<b>p-value</b>
Intercept	<b>0.4520**</b>	0.04635	3.715	0.0098	<b>0.45341**</b>	0.0345	4.098	0.003	0.2710	0.56341	0.935	0.345
IVI	<b>0.1123**</b>	0.94524	4.945	0.005	0.4321	0.9834	0.9341	0.453	0.4581	0.76531	0.845	0.301
Carotenoids	0.15797	0.46552	0.945	0.235	-0.5761	0.8351	-0.974	0.645	0.3312	0.86754	0.424	0.452
Proline	<b>0.09453**</b>	0.00232	5.734	0.000	<b>0.5643**</b>	0.0845	5.735	0.0031	<b>0.3412**</b>	0.09843	3.872	0.003
Chlorophyll-a	<b>-0.0945*</b>	0.00342	-3.098	0.031	<b>-0.6451*</b>	0.0099	-2.971	0.021	-0.3519	0.46341	-0.452	0.471
Chlorophyll-b	<b>-0.0576**</b>	0.09130	-4.452	0.0082	<b>-0.034*</b>	0.94531	-1.986	0.031	-0.7451	0.56341	-0.741	0.234
<b>Random effect</b>	<b>Variance</b>	<b>St. dev</b>			<b>Variance</b>	<b>St. dev</b>			<b>Variance</b>	<b>St. dev</b>		
All Zones	0.1451	0.74534			0.241	0.3710			0.3710	1.13203		
Zone (LPZ-HPZ)	0.0413	0.13367			0.0915	0.2345			0.11444	0.56628		
Zone (HPZ-MPZ)	0.0354	0.34192			0.0556	0.74614			0.09634	0.463321		
Zone (MPZ-LPZ)	0.1089	0.36718			0.1445	0.75674			0.25743	0.96345		
Residual	0.4810	0.1345			0.7513	0.3410			0.8140	0.2413		











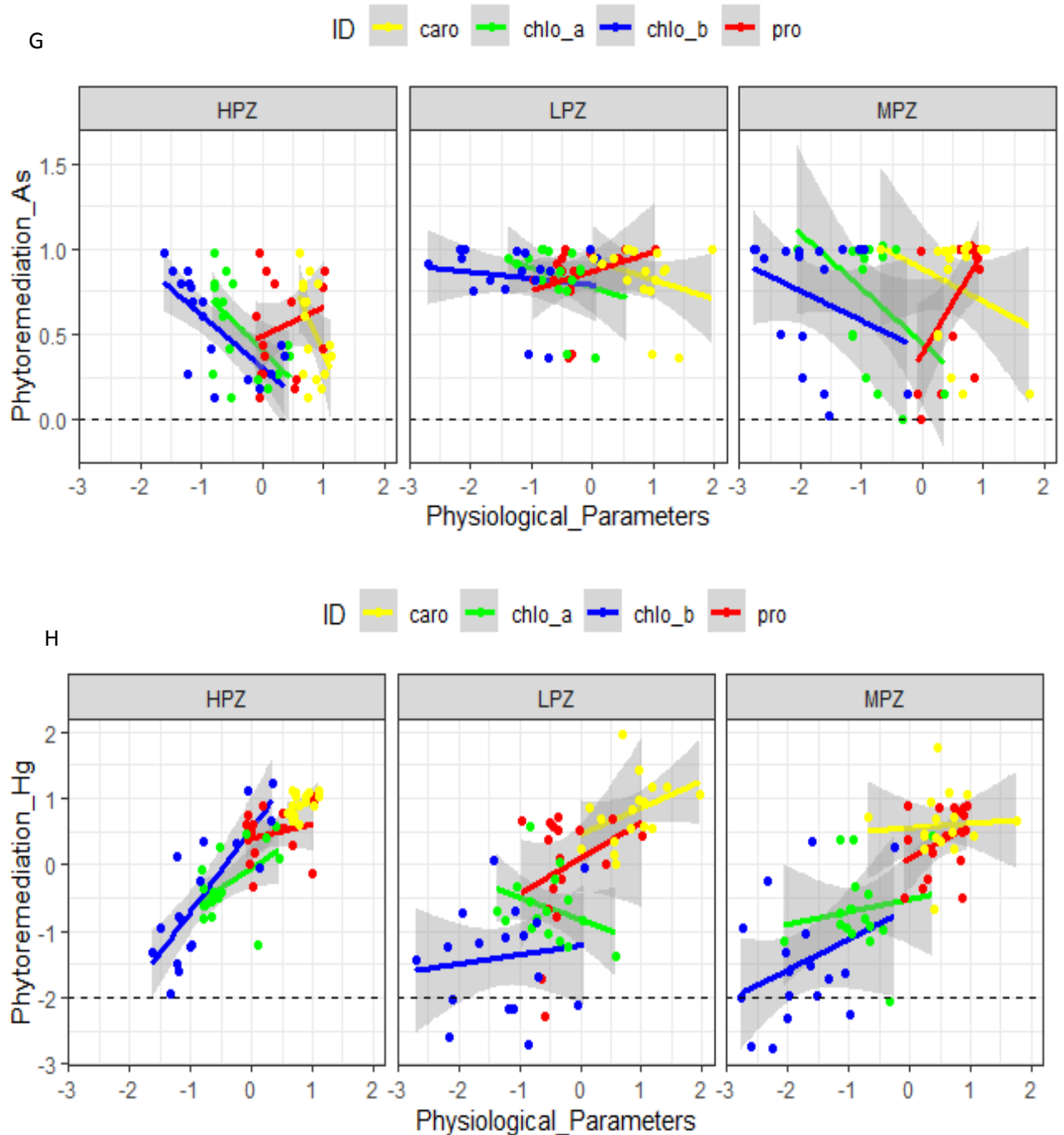


Figure 5. 6 (A-H) Graphical representation of GLMM; Phytoremediation Impact of various heavy metals on Plants's physiological parameters in three distinct polluted zones, yellow showed carotenoid, green Chlorophyll a, blue Chlorophyll b, and red line the proline content.



## 5.4 Discussion

Environmental pollution has increased drastically in recent decades, posing considerable ecological and human health threats. The study promotes phytoremediation, a low-cost, plant-based technique, as a successful method for mitigating these consequences. Our research observed the phytoremediation capacity of several indicator plants, including *Achyranthes aspera*, *Arundo donax*, *Calotropis procera*, *Cannabis sativa*, and others, by calculating their Bioaccumulation Coefficient, Translocation Factor, and Biological Concentration Factor. These plants showed a remarkable ability to remove heavy metals such as Zn, Cu, Ni, Cd, Cr, Ar, Pb, and Hg from environments polluted by industrial wastewater. Our findings have a strong link with previous literature, particularly the work of (Sun, Liao et al. 2014) on phytoremediation in the industrial zones of South-Central China. Our findings add to our understanding of the efficacy of plant species in the remediation of heavy metals such as Pb, Cd, Zn, Ni, and As in polluted soil, including *Arundo donax*, *Calotropis procera*, *Cannabis sativa*, and *Parthenium hysterophorous*. This aligns with findings from various studies (Al-Yemni, Sher, El-Sheikh, Eid, & Essays, 2011; D'Souza, Varun, Masih, & Paul, 2010; Papazoglou, Serelis, & Bouranis, 2007). A pivotal insight from our research is the exceptional ability of certain weedy species, such as *Calotropis procera*, *Cyperus rotundus*, *Datura stramonium*, and *Parthenium hysterophorous*, which spontaneously flourish in industrial zones and demonstrate a high capacity for heavy metal accumulation. This finding offers a practical and sustainable approach to environmental management, resonating with the research of (N. Kumar et al., 2013) and (Lum, Ngwa, Chikoye, & Suh, 2014) and suggests that promoting these species via direct cultivation or as part of wild plant communities, could be a strategic move for improved environmental remediation.

The current research demonstrates that the Bioconcentration Factor (BCF), Translocation Factor (TF), and Bioaccumulation Coefficient (BAC) values for all selected plants suggest these species are effective in phytostabilization and phytoextraction in ecosystems polluted by industrial wastewater. This aligns with (R. S. Kumar & Thambavani, 2012) study on roadside vegetation in polluted areas, which identified species like *Pongamia pinnata*, *Polyalthia longifolia*, *Azadirachta indica*, and *Ficus religiosa* as effective in pollution remediation. Similarly, (Noor et al., 2015) assessed the impact of pollution on vegetation near industries using the Air Pollution Tolerance Index and Anticipated Performance Index, noting considerable effects. Similarly, (Malik, Husain, et al., 2010) also studied how different plant species accumulate various metals, concluding that factors like



plant species, soil conditions, climate, transfer mechanisms, sequestration, root system type, response to elements, and seasonal cycles significantly influence the heavy metal accumulation and bioavailability. (H. Liu et al., 2018) higher soil pH significantly reduces heavy metal concentration and leaching due to decreased solubility in less acidic soils, thus lowering the uptake and translocation of these metals into plant tissues.

In exploring the research question, "How do physiological adaptive responses in indicator plants combine to affect their resilience to metal toxicity in polluted ecosystems?" our study delved into the relationship between proline accumulation and the health of photosynthetic pigments in the context of heavy metal stress. Our findings revealed a significant positive correlation between increased proline levels and tolerance to metal toxicity in 17 indicator plant species, more pronounced in a high polluted zone (HPZ) than in a low polluted zone (LPZ). This finding is consistent with proline acting as an effective defense mechanism against metal toxicity. This observation aligns with (Heuer & stress, 2010) and (S. Qadir et al., 2014), who have similarly reported the role of proline in enhancing plant tolerance to mental stress. Simultaneously, our study observed a consistent decrease in chlorophyll-a, chlorophyll-b, and total carotenoids across all indicator species involved in heavy metal remediation, indicating a negative impact of pollution on these essential photosynthetic pigments. This finding is in line with the observations made by (Aggarwal et al., 2012) and (Zengin, Munzuroglu, & Science, 2006), who noted similar reductions in photosynthetic pigments under metal stress. The decline in these pigments is crucial as it reflects the environmental stress on plant systems and underscores the importance of developing effective phytoremediation strategies. When plants are exposed to different types of pollution, their photosynthetic pigments, specifically carotenoids and chlorophyll, are reduced in concentration. This exposure also impacts pedicle length, overall plant yield, seed germination, and the number of inflorescences, as observed by (Nithamathi & Indira, 2005). Pollution and heavy metal stress can decrease enzyme activity crucial for chlorophyll synthesis, resulting in diminished photosynthetic activity, as noted by (De Filippis & Pallaghy, 1976). Furthermore, (Schickler & Caspi, 1999) reported disruptions in membrane integrity, and (Kastori, Plesničar, Sakač, Panković, & Arsenijević- Maksimović, 1998) found that pollution leads to metal ion exchange in chlorophyll molecules and a decrease in leaf chlorophyll concentration.

Significant reductions in chlorophyll content have been observed during drought stress in various plant species, including *Catharanthus roseus*, *Gossypium hirsutum*, *Helianthus annuus*, and *Vaccinium myrtillus*, as documented in studies by (Jaleel, Sankar, Sridharan, &



Panneerselvam, 2008; Kiani, Maury, Sarrafi, & Grieu, 2008; Lawlor & Cornic, 2002; Tahkokorpi, Taulavuori, Laine, & Taulavuori, 2007). (Lawson, Lefebvre, Baker, Morison, & Raines, 2008) a decrease in relative water content and leaf water capacity correlates with a lower foliar photosynthetic rate in higher plants. This reduction may result from metabolic impairments or stomatal closure.

Furthermore, our findings help to clarify the intricate link between proline production and plant stress response. While an increase in proline production has been linked to improved resilience and recovery from heavy metal-induced physiological stress (Gajewska & Skłodowska, 2008) (Soares et al., 2016), It does not always result in a decrease in growth inhibition or oxidative damage in plants. These complexities in reaction have been addressed in investigations by (Handique & Handique, 2009) and (Rizwan et al., 2017), proving that proline accumulation could predict a plant's vulnerability to stress.

Proline plays a crucial role in plant physiology, acting as a protectant for certain enzymes against denaturation, serving as a nitrogen and carbon source, stabilizing protein synthesis, regulating cytosolic acidity, and scavenging hydroxyl radicals, as identified by (Mateos et al., 2020). Plants increase their proline concentration as a defense mechanism to cope with environmental stress, enhancing survival (Akshita, Nandini, Sumedha, & Trishang, 2018). Indeed, exposure to pollutants can reduce the number of photosynthetic pigments in plants, as observed in various studies. Despite this challenge, plants adapt through various physiological responses that enable them to utilize resources more efficiently under stressful conditions. These adaptations are critical for plants' growth and survival in polluted environments (Bukhari et al., 2021; Ullah, Al-Busaidi, Al-Sadi, & Farooq, 2022). (Zouari et al., 2016) they explored the beneficial role of exogenous proline in plants under cadmium heavy metal stress. Adding proline externally can enhance a plant's antioxidant defense mechanisms and improve mineral uptake. This, in turn, helps reduce the oxidative damage caused by cadmium exposure. Proline also increases photosynthetic activities and mineral nutrition under salt stress in *Olea europaea* (olive tree), as noted by (S. Lal, Ahmed, Srivastava, & Singh, 2015). Similarly, (J. Xu, Yin, & Li, 2009) reported that exogenous proline enhanced heavy metal tolerance in *Solanum nigrum* by improving antioxidant enzyme activities. (M. Singh, Singh, Dubey, & Prasad, 2015) proline application reduces the toxic effects of arsenate in *Solanum melongena* by decreasing arsenate accumulation and oxidative stress.

(S. Hayat et al., 2012) and (Shahid et al., 2014) observed that plants exposed to adverse environmental conditions accumulate proline for various functions, such as maintaining



osmotic balance, scavenging reactive oxygen species, and stabilizing cell membranes. The beneficial effects of proline can be direct, like increasing photosynthetic rate and mineral nutrition, or indirect, such as enhancing disease tolerance. (Shahid et al., 2014) also reported that exogenous proline concentration increased fresh plant weight in *Pisum sativum* by enhancing CO<sub>2</sub> absorption for photosynthesis. (Dawood, Taie, Nassar, Abdelhamid, & Schmidhalter, 2014) linked the protective effect of exogenous proline to improved mineral uptake. Our findings provide a more detailed understanding of the integrated physiological adaptation responses in plants to metal toxicity. The dual role of proline as a defense mechanism and a possible signal of stress susceptibility, along with a decrease in photosynthetic pigments, highlights the complex nature of plant adaptive responses in contaminated environments. These findings are critical for informing future research and the development of more effective phytoremediation technologies for dealing with metal toxicity. The outcomes of our study provide vital insights for solving environmental concerns in industrial areas afflicted by heavy metal contamination. Our findings highlight the potential efficacy of some plants in phytoremediation but also imply that a broader, more holistic approach may be advantageous. This would include developing a diverse range of flora, either naturally or through intentional planting and sowing, which might increase the effectiveness of removing pollutants from these ecosystems. Moreover, some recent studies have supported the idea that removing these plants regularly by harvesting could effectively extract accumulated chemicals. (Ji, Sun, Song, Ackland, & Liu, 2011; Muthusaravanan et al., 2020).

Our findings reveal that several plants, including *Achyranthes aspera*, *Calotropis procera*, *Parthenium hysterophorus*, and *Cannabis sativa*, have outstanding phytoremediation abilities. These species, already known for their incredible environmental detoxification capacities, could play an important role in developing a more efficient and diversified phytoremediation strategy. (Ejaz et al., 2022a; Kaur et al., 2023; Linger, Müssig, Fischer, & Kobert, 2002; Sarfraz et al., 2022a; Usman, Al Jabri, Abu-Dieyeh, & Alsafran, 2020). The different benefits of each plant could be used to boost the overall effectiveness of phytoremediation approaches.

This research establishes a basis for further investigations, primarily aimed at unravelling the molecular processes behind phytoremediation, examining the detailed function of proline in plant stress responses, and assessing the broader ecological effects of employing phytoremediation techniques. Future studies could also investigate the genetic



diversity in phytoremediation capabilities across various plant species, aiming to develop more precise and effective environmental remediation approaches through phytoremediation.

## 5.5 Conclusion

The escalating pace of industrialization and increasing human populations call for urgent protective measures for our environment against excessive exploitation and pollution. Developing nations, in particular, face challenges in upgrading their industrial facilities, many of which are not eco-friendly. This inadequacy in addressing metal pollutants has pronounced adverse impacts on local ecosystems, leading to detrimental physical and chemical alterations and their direct negative impacts on human health. Overall, such metal pollution impairs ecosystem integrity and affects the health and livelihoods of local communities, flora and fauna, causing diseases, compromising food safety, and resulting in habitat destruction. In light of these challenges, our research has identified pollution-tolerant plant species with high phytoremediation potential, notably *Achyranthes aspera*, *Arundo donax*, *Calotropis procera*, *Cannabis sativa*, *Chenopodium album*, *Coronopus didymus*, *Cynodon dactylon*, *Dysphania ambrosioides*, *Eclipta alba*, *Ficus carica*, *Koeleria macrantha*, *Malva neglecta*, *Ricinus communis*, *Parthenium hysterophorus*, *Persicaria glabra*, *Saccharum bengalensis*, and *Ziziphus nummularia*. These plants exhibit remarkable efficacy in the phyto-extraction and phyto-stabilization of heavy metals from industrially polluted ecosystems. Promoting these species offers a potential strategy for enhancing environmental sustainability and resilience. Additionally, our findings reveal a consistent increase in proline osmoprotectant and a decrease in chlorophyll-a, chlorophyll-b, and total carotenoids in response to elevated metal pollution. These physiological adjustments suggest an adaptive mechanism for coping with pollutant stress. This provides mechanistic support for promoting these species in and around polluted ecosystems via cultivation or otherwise to aid in environmental clean-up efforts. This study's methodologies and analytical tools provide a replicable model for ecosystem remediation efforts globally. By adopting such approaches, we can better monitor and implement environmental strategies that foster sustainable development, ensuring that the benefits extend beyond the immediate surroundings to global ecological systems.





Contents lists available at ScienceDirect

Journal of Environmental Management

journal homepage: [www.elsevier.com/locate/jenvman](http://www.elsevier.com/locate/jenvman)



Research article

## Elucidating the phytoremediation potentials and ecophysiological mechanisms of indicator plants in the industrial polluted region

Ujala Ejaz<sup>a,b</sup>, Shujaul Mulk Khan<sup>a,c,\*</sup>, Noreen Khalid<sup>d,e</sup>, Sadia Jehangir<sup>a,b</sup>,  
Shah Fahad Ali Shah<sup>f</sup>, Jens-Christian Svenning<sup>b,\*\*</sup>

<sup>a</sup> Department of Plant Sciences, Quaid-i-Azam University, Islamabad, 45320, Pakistan

<sup>b</sup> Center for Ecological Dynamics in a Novel Biosphere (ECONOVO), Department of Biology, Aarhus University, Ny Munkegade 114, DK-8000, Aarhus C, Denmark

<sup>c</sup> Member Pakistan Academy of Sciences, Pakistan

<sup>d</sup> Department of Botany, Government College Women University, Sialkot, Pakistan

<sup>e</sup> Faculty of Health and Life Sciences, INTI International University, 71800, Nilai, Negeri Sembilan, Malaysia

<sup>f</sup> School of Economics and Management, Yanshan University, Hebei Province, 066004, China

### ARTICLE INFO

#### Keywords:

Heavy-metals  
Industrial pollution  
Indicator plants  
Proline  
Global south  
Bioremediation

### ABSTRACT

The integrity of natural ecosystems, particularly in the Global South, is increasingly compromised by industrial contaminants. Our study examines the growth of plant species adapted to ecosystems impacted by heavy metal pollution, specifically focusing on their phytoremediation capabilities and tolerance to contaminants. The potential of pollution-tolerant species was evaluated in the industrial subtropical wetland of Sialkot, Pakistan. Employing quantitative ecological methods, data on vegetation, phytosociological attributes, and soil properties were gathered from 450 plots across different pollution gradients. The study pinpointed 17 key indicator species tolerating high heavy metal pollution out of 182 surveyed, using a combination of Indicator Species Analysis (ISA) and the Importance Value Index (IVI). These species demonstrated diverse capacities to extract, stabilize, and accumulate heavy metals (Cr, Zn, Cu, As, Cd, Ni, Hg, and Pb) across varying pollution zones. Notably, *Cannabis sativa* demonstrated substantial phytoextraction of Zn and Cd, with concentrations reaching 1977.25 µg/g and 1362.78 µg/g, respectively. *Arundo donax* showed marked hyperaccumulation of Cd, peaking at 410.531 µg/g. *Achyranthes aspera* was remarkable for its extraction and accumulation of Ni and Cu, with concentrations of 242.412 µg/g and 77.2997 µg/g, respectively. Physiological changes, such as increased proline levels in *Cannabis sativa* and *Achyranthes aspera* reaching 39.041 µg/g and 27.523 µg/g under high metal concentrations, indicated adaptation to metal stress. Declines in chlorophyll and carotenoid levels were also observed as metal contamination increased, with up to 35% reductions in some species. These findings underscore the potential efficacy of selected plant species in phytoremediation and highlight the importance of physiological responses in their tolerance to metals, providing valuable information for targeted remediation strategies in polluted ecosystems and improving environmental management and sustainable practices.



# Monitoring Land Use Dynamics: Remote Sensing and NDVI Analysis for Assessing Urban-Industrial Expansion and Ecological Impacts in Aik-Stream

## 6.1 Introduction

Land use patterns and ecosystems have substantially changed due to the increase of built-up areas caused by urbanization and industrialization (H. Jiang et al., 2022). There is a significant shift from rural to urban environments due to concrete, asphalt, and metal infrastructure around these areas (Kuussaari et al., 2021). Globally, 54% of people lived in urban areas in 2014; by 2050, that number is expected to climb to 66%. Although these revolutions are happening throughout the globe, the rate of fluctuation depends on region and place (Hoornweg & Pope, 2017). In North America, 82% of the population lives in urban and industrial areas, compared to 40% and 48% in Africa and Asia (Madlener & Sunak, 2011). African and Asian regions will continue to expand fastest, with urban and industrial populations predicted to reach 56% and 64% by 2050 (C. Zhong et al., 2023). Urban and industrial development is growing slower, even in more developed regions like North America (Rondinelli, Johnson Jr, & Kasarda, 1998). This increase in the growth of built-up areas emphasizes the importance of implementing comprehensive strategies for monitoring, managing, and analyzing the numerous difficulties posed by urbanization and industry (X. Wang, 2020). This development adversely affects agriculture, forestry, water quality, and environmental conditions (Hollis, 1990). In order to ensure the long-term sustainability of urban development and safeguard natural resources for generations, addressing the environmental consequences of expanding urban areas is crucial (de Vries, 2021).

### 6.1.1 Land Use and Land Cover Change (LULC)

Land Use and Land Cover (LULC) changes through urbanization and industrialization are the significant ways human activity has altered the planet (Y. Wu, Li, & Yu, 2016). Landscape mapping, known as LULC mapping, examines these transformations by categorizing the environment into segments such as wetlands, urban regions, forests, and bare land areas. To precisely evaluate developed areas, vegetation cover, water resources, and land growth, it is crucial to outline each LULC category (Yin et al., 2011). The USGS National Land Cover Database classifies urban areas according to the geographic density of developed land cover (C. H. Homer, Fry, & Barnes, 2012). Remote sensing has proven to be an asset in



monitoring growth, leveraging the diverse visual and spectral characteristics of various land use and land cover types (C. Zhang & Li, 2022). The versatility of sensing in capturing timeframes and spatial dimensions has proven to be extremely useful for analyzing urban and industrial expansion, resulting in the creation and utilization of various image processing methods to improve the precision of land use and land cover research (Aqil Tariq & Mumtaz, 2023). Various methods for classification, like density slicing, the Normalized Difference Vegetation Index (NDVI) and the Index-based Built-Up Index (IBI), are some examples of these techniques (Sinha, Verma, & Ayele, 2016).

### ***6.1.2 Urban-industrial Development and Riparian Areas***

Areas directly adjacent to bodies of water, known as riparian zones, extend to regions up to 100 meters away and beyond. These areas are crucial to protecting water quality by acting as barriers that capture pollutants from areas before they enter aquatic environments (Ferreira, Aguiar, & Nogueira, 2005). The plants along the riverbanks play a role in purifying sediment and nutrients from above ground and underground water, improving the water quality (Odum, 1979). The effects also shape the plant populations and different water quality measures. Nevertheless, alterations in land use and land cover (LULC) threaten these areas by disturbing their harmony and changing their scenery (Muriithi, 2016). The study indicates that it is crucial to have vegetated buffers at least 30 meters wide to manage non-point source pollution effectively (Ferreira et al., 2005). The closeness of surfaces to water sources and streamside buffers is recognized to impact the quality of water (K. Li et al., 2018). Therefore, more attention is being focused on studying and developing policies to protect areas. These areas are valued for their ability to provide insights into the quality of stream water compared to looking at the land cover of a watershed. In cases where more than 10% of surfaces are found in riparian areas it could indicate the beginning of stream deterioration (Schiff & Benoit, 2007).

### ***6.1.3 Mapping urban-industrial development***

Historically, land use and land cover (LULC) change investigations were based on in situ surveys and aerial photography (Paudel et al., 2016). Though these approaches provide thorough information, their application is limited by resource requirements' time and space constraints. In contrast, remote sensing and Geographic Information Systems (GIS) provide methods for mapping and studying historical as well as current Land Use and Land Cover (LULC) alterations across vast regions, exceeding conventional methods (Mohamed &



Plante, 2002). Remote sensing explicitly allows data to be gathered in certain areas, improving the effectiveness of analyzing changes over time (Sabins Jr & Ellis, 2020).

The development of remote sensing and GIS technologies has transformed how we gather and study Land Use and Land Cover (LULC) changes and provides a budget-comprehensive approach to exploring urban growth (Dhanaraj & Angadi, 2022). Satellite optical sensors are mainly utilized for collecting land use and land cover (LULC) data, as they can differentiate between types of LULC by analyzing the responses (LAKEW, 2020). The collaboration between USGS and NASA in the Landsat missions, which began in 1972 and includes sensor improvements, showcases the importance of sensing in land use and land cover research (Almeida, Teodoro, & Gonçalves, 2021). The Landsat information plays a role in creating the National Land Cover Database (NLCD), a resource that is updated every five years by the Multi-Resolution Land Characteristics Consortium (MRLC) to produce a map of land cover across the entire country (C. Homer et al., 2015).

Optical sensors on satellites, such as the ones used in Landsat missions, play a role in mapping the growth of cities and changes in land use and land cover (ED Chaves, CA Picoli, & D. Sanches, 2020). Lately, for research development, there has been an increased fascination with sensor technologies, like synthetic aperture radar (SAR) and light detection and range (LiDAR). Enhancing the precision and comprehensiveness of information gathered in studies is achieved by combining data from active sensors (García-Pardo, Moreno-Rangel, Domínguez-Amarillo, & García-Chávez, 2022). Commonly utilized methods, like the Normalized Difference Vegetation Index (NDVI), identify the vegetation cover of the study area (Zheng, Tang, & Wang, 2021). Furthermore, supervised classification algorithms play a role in identifying patterns for land use and land cover (LULC) classification. Despite the difficulties, it can be challenging to differentiate between land use and land cover categories (Kathar, Nagne, Awate, & Bhosle, 2023).

#### ***6.1.4 Analysis Techniques***

Urbanization, industrialization development dynamics, and environmental and social consequences rely comprehensively on land use and land cover (LULC) transformations (Das & Angadi, 2020). Remote sensing and Geographic Information Systems (GIS) have gradually enhanced and often surpassed direct surveys and aerial imagery analysis in recent decades (Ozenen Kavlak, Cabuk, & Cetin, 2021). These innovations have significantly affected urban expansion, environmental change, and human-nature interactions. Accurate mapping and



understanding are essential to plan and manage built-up areas sustainably and effectively (Tomaszewski, 2020). Through the emergence of machine learning algorithms, LULC change research has become more sophisticated (J. Wang, Bretz, Dewan, & Delavar, 2022). These algorithms clarify the complex and nonlinear relationships that characterize built-up areas and the evolution of vegetation (Bindajam, Mallick, Talukdar, Islam, & Alqadhi, 2021). Machine learning can be applied to dissect historical urbanization trends and forecast future trends more accurately. The rapid growth of built-up areas caused by urbanization and industrialization will significantly alter land use and cover. In urban and industrial areas, riparian areas are negatively affected by impervious surfaces and changes in land cover composition. This reduces water quality and ecological function (Cunha, FARIA, NASCIMENTO, Silva, & CUNHA, 2021). Urbanization and industrialization's spatial and temporal dynamics can be mapped and analyzed using remote sensing and GIS techniques (Murayama, Simwanda, & Ranagalage, 2021). Machine learning algorithms combined with remote sensing data will improve the ability to predict and understand complex urbanization and vegetation growth patterns (Kalantar et al., 2020).

LULC studies take place in areas where remote sensing, GIS, and machine learning can be used to support them. This chapter examines the cutting-edge applications of these tools. This chapter emphasizes the critical role of technological advancements in augmenting our understanding of urban-industrial development trajectory. Furthermore, it illuminates the inherent challenges and potential pathways for achieving sustainable urban-industrial development, marking an essential chapter in the broader history of environmental and urban planning (Kim, de Leeuw, Harris-Roxas, & Sainsbury, 2023).

#### ***6.1.5 Statement of the Problem and Scope of the Study***

Several urbanized sub-watersheds have been centered in Sialkot, Pakistan, over the last few decades. Rapid urban and industrial expansion highlights the importance of understanding how urban and industrial growth affects local water bodies, as impacts on surface water quality pose substantial challenges to sustainable development. Aik-Stream flows through and around Sialkot, and this research aims to assess the impact of industrial expansion on its water quality using satellite imagery and geospatial data. In this study, changes in vegetation will be examined as an indicator of the extent to which urbanization and industrialization have altered ecosystems and land use patterns. As another objective, we aim to analyze the spatial evolution of built-up areas in relation to land use patterns, applying different machine learning techniques to decipher complex development dynamics. Further,



the study proposes effective strategies for sustainably managing urban-industrial growth, considering social, economic, and environmental factors. A pioneering effort has been undertaken in this study to evaluate the spatial relationships between local water resources and the growth of urban-industrial activity in the region. The overall landscape dynamics of the wider Sialkot area are examined by using remote sensing data and GIS analysis.

## **6.2. Material and Methodology**

### ***6.2.1 Research Design***

This study examined how land use and land cover changed within a designated area over the past two decades due to changing land use and land cover patterns. The satellite imagery used for this purpose came from the Landsat 5 Thematic Mapper (TM) and the Landsat 8 Operational Land Imager (OLI). Historical coverage is abundant on these satellites, and they provide high-resolution images, so they are ideal for GIS.

A Landsat image was the first piece of information that needed to be acquired and prepared as a part of our research process. The radiometric and atmospheric anomalies were corrected to ensure the imagery was reliable and consistent for in-depth analysis. Supervised image classification techniques were used to categorize the landscape accurately using LCLU classification. In order to map the diverse land uses in the area and understand how they have evolved, it was crucial to complete this phase. Prior to focusing on the ecological aspects of the study, the validation of LCLU classification results was the priority. To confirm the accuracy and reliability of our initial classifications, we conducted a thorough accuracy assessment against ground truth data. In addition to being a procedural requirement, the basis for our analysis needed to be robust and accurate. Once our LCLU classification was established with high accuracy, we conducted the NDVI analysis. After assessing the health of the vegetation in the study, we evaluated its distribution and health. NDVI (Normalized Difference Vegetation Index) was calculated to determine the ecological consequences of LCLU changes. A better understanding of the environmental dynamics over the study period was achieved by identifying changes in vegetation patterns. According to Figure 6.1, the imagery and mythological pattern were processed and analyzed.



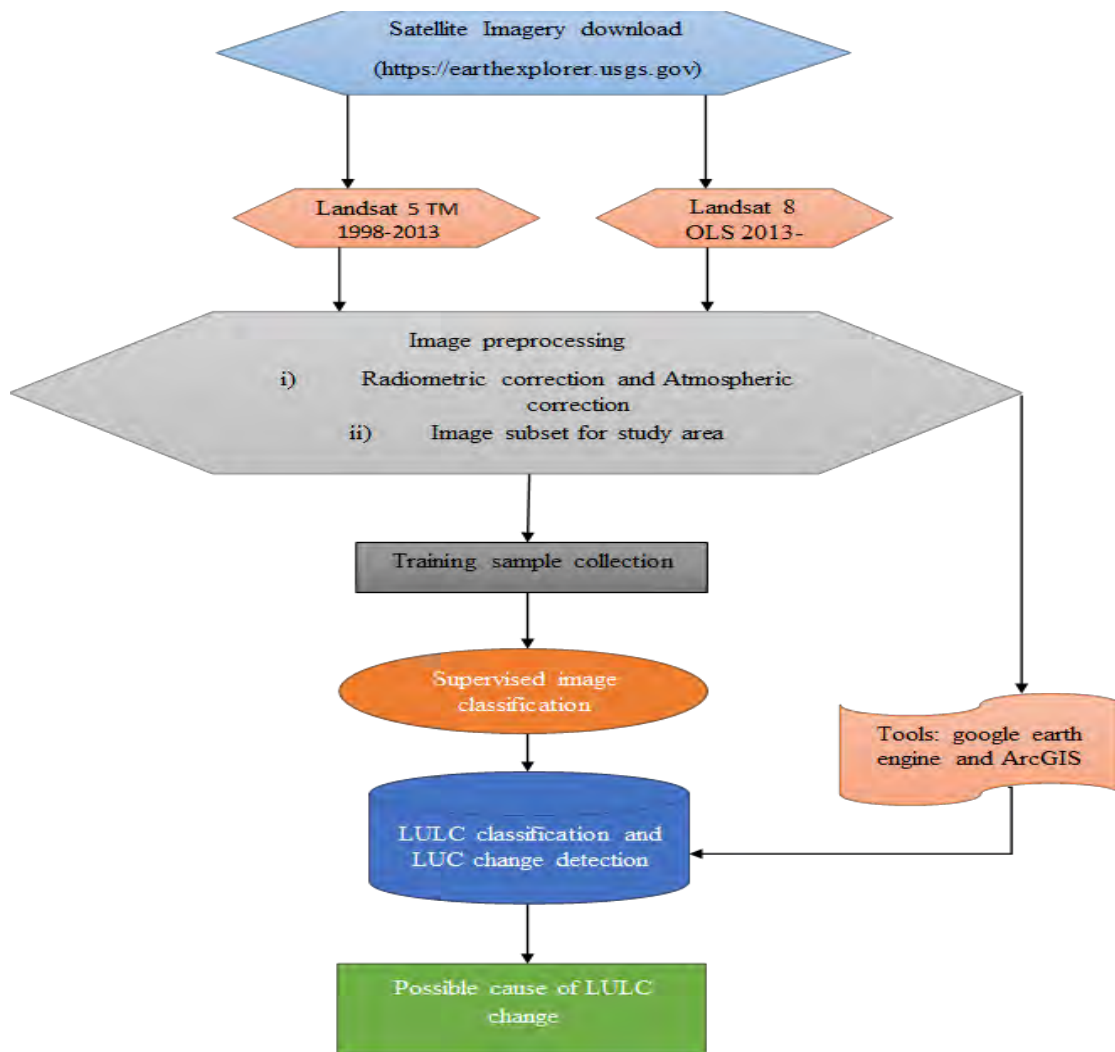


Figure 6.1: NDVI, land use, and land cover data collection, processing, and methodological scheme.

### 6.2.2 Data Acquisition

Satellite images used in this study were taken from the USGS Earth Explorer (<http://earthexplorer.usgs.gov>). The study area falls within a specific section of imagery captured by the Landsat satellite, identified as path 172 and row 62 within the Worldwide Reference System (WRS). This study utilized Landsat 4-5 Thematic Mapper (TM) and Landsat 8 Operational Land Image (OLI) scenes for almost 25 years (1998-2023). Information about the Landsat satellites and their spectral bands can be found at <https://landsat.gov/what-are-band-designationslandsat-satellites>. Obtaining satellite images without clouds is a big challenge in tropical areas. The images we collected have been adjusted for terrain at level 1, making them ideal for studying individual pixel-level changes over time. Both satellite images in the Red Green Blue (RGB) and Near Infra-Red (NIR) spectrums have collected images were provided in Geo Tiff format and resampled to 30 m



with Universal Transverse Mercator (UTM) World Geodetic System (WGS) 84 projection using the cubic convolution method (ED Chaves et al., 2020).

### **6.2.3 Image preprocessing**

To analyze data effectively, it is essential to ensure that the images are both geometrically and radiometrically accurate. Before conducting any analysis, image preprocessing should be performed to provide a dataset suitable for extracting spatial information (C. Yang, Jia, Dong, Zhao, & Zhao, 2024). This step helps eliminate geometric and radiometric distortions caused by objects' shape, color, and the atmosphere's effects on satellite images. Image processing is prepared for analysis by applying various steps: (i) converting them to radiance values, (ii) adjusting atmospheric conditions, and (iii) correcting distortion in their geometry. The image processing and classification were performed using ArcGIS.

### **6.2.4 Image classification and training sample collection**

The study area was categorized into four land use and land cover classes during a specific period. These categories/classes include water, bare land, built-up areas, and vegetation, encompassing all plants and trees. Historical high-resolution imagery from Google Earth (GE)TM was used to gather sample training data for image classification to identify different land cover types. A minimum of 100 training samples for each class were collected to ensure higher accuracy. Google Earth (GE) provided reference images for 1998; the current image (2023) was used for accurate classification. The Landsat series of satellites have wavelength bands and spectral characteristics that are very similar to each other, which simplifies the process of collecting samples. Random Forest (RF) supervised classification method was used to perform the image classification. This method involves three main stages: training, class allocation, and testing. During the training stage, the regions of interest (ROIs) for various land use and land cover (LULC) categories were defined in ENVI 6.1, utilizing historical imagery from Google Earth (GE).

Table 6.1 Classification of the satellite images for the study area from 1998 to 2023.

Classification	Description
Vegetation	Including all agriculture and natural forest
Built-up area	Including all residential, industrial, and transportation
ater	All water bodies
Bare land	Including all those land areas influenced by human



### 6.2.5 Normalized Difference Vegetation Index (NDVI)

The Normalized Difference Vegetation Index (NDVI) was initially created to map areas with green vegetation (S. Huang, Tang, Hupy, Wang, & Shao, 2021). This index relies on the significant difference in how red wavelengths are absorbed compared to near-infrared wavelengths by healthy vegetation. Healthy green plants absorb much light in the red part of the Electromagnetic Spectrum (EMS) but reflect much of the near-infrared (NIR) part. This behavior is mainly due to the plant's photosynthesis process. The contrast in absorption across the electromagnetic spectrum (EMS) means that regions with healthy, green plants typically show high positive values in the normalized difference vegetation index (NDVI). Additionally, the Normalized Difference Vegetation Index (NDVI) can help map areas without vegetation, such as water bodies and impervious surfaces (IS), because each land cover type interacts differently with red and NIR wavelengths, allowing for their differentiation. The formula for this index, as shown in equation 6.1, utilizes both the near-infrared (NIR) and red (R) parts of the Electromagnetic Spectrum (EMS):

$$NDVI = \frac{(\rho_{NIR} - \rho_R)}{(\rho_{NIR} + \rho_R)} \dots \dots \dots (6.1)$$

Where  $\rho_{NIR}$  The surface reflectance is the value of NIR and  $\rho_R$  is the surface reflectance value of R.

Equation 2 shows the formula for calculating the Normalized Difference Vegetation Index (NDVI) using Landsat 5 Thematic Mapper (TM) data, while equation 6.3 demonstrates the NDVI formula specific to Landsat 8 Operational Land Imager (OLI).

$$Landsat\ 5\ TM\ NDVI = \frac{(\rho_{B4} - \rho_{B3})}{(\rho_{B4} + \rho_{B3})} \dots \dots \dots (6.2)$$

Whereas  $\rho_{B4}$  and  $\rho_{B3}$  is represents the reflective value of NIR and R, respectively.

$$Landsat\ 8\ OLI\ NDVI = \frac{(\rho_{B5} - \rho_{B4})}{(\rho_{B5} + \rho_{B4})} \dots \dots \dots (6.3)$$

Whereas  $\rho_{B5}$  and  $\rho_{B4}$  represents the reflective value of NIR and R, respectively.

### 6.2.6 Accuracy Assessment

The accuracy assessment of the result of the LULC dataset was conducted using two different techniques (1) zonal statistics for every 5 year gap (2) Kappa statistics (Gaur, Mittal,



Bandyopadhyay, Holman, & Singh, 2020). Based on this statistic, LULC models from 1998-2003, 2003-2008, 2008-2013, 2013-2018, and 2018-2023 were assessed for accuracy. A random sample of 100 points was selected for every zonal accuracy assessment.

### 6.2.7 Kappa Statistics

Cohen's Kappa-coefficient (k) and confusion matrix (CM) were calculated for each date. The Kappa statistic is commonly used in remote sensing to evaluate classification models. As a result, it determines inter-classifier agreement and mitigates biases. As a result of this statistic, the effectiveness of the IS classification model can be evaluated by calculating the accuracy of the User (Type I error or false positive) and Producer (Type II error or false negative) for each class, as well as the Kappa coefficient of agreement. The producer's accuracy determines the model's omission error, which shows how much of a specific class was incorrectly classified. In Equation 6.4, the accuracy of the Producer is calculated.

$$\text{producers Accuracy} = \frac{C_t}{C_t + O_c} \dots \dots \dots (6.4)$$

The CT represents the number of class Cs correctly classified, and the OC represents the sum of the other classes classified as C.

Pixels correctly classified (PCC) are determined by dividing the model's total accuracy by its total classification. This calculation helps gauge the model's accuracy in recognizing the correct pixels.

$$\text{User Accuracy} = \frac{C_t}{C_t + C_E} \dots \dots \dots (6.5)$$

The Kappa coefficient, denoted as K, indicates how well a classification model performs compared to random chance, considering that some correct classifications may occur by coincidence. It measures the agreement between the model's predictions and actual classifications. The formula to calculate the Kappa coefficient is provided by Equation (6).

$$K = \frac{p_0 - p_E}{1 + p_E} \dots \dots \dots (6.6)$$

PO represents how much agreement exists in the observed classification. PE, on the other hand, signifies the agreement we would expect by chance alone. When we have a complete agreement, the value is 1, indicating the highest level of agreement possible.



### 6.3 Results

The study utilized Landsat imagery from 1998, 2003, 2008, 2013, 2018, and 2023 to categorize Sialkot's land use and cover types. These categories included vegetation, water bodies, urban areas, and bare land. The research undertaken from 1998 to 2023 within the Sialkot area of Pakistan, with a focus on the surroundings of the Aik stream, exposes significant transformations in the patterns of land use and land cover over forty years. These transformations are comprehensively catalogued in (Table 6.2, which provides an exhaustive account of the shifting landscape dynamics in this area.

The study highlights a notable increase in urbanized and industrialized zones, with the built-up area expanding by 15.84 km<sup>2</sup> from 10.07 km<sup>2</sup> in 1998 to 25.91 km<sup>2</sup> by 2023. This growth reflects considerable urban expansion and industrial development, significantly altering nearby natural and semi-natural environments. Conversely, the vegetation class suffered a marked decrease, shrinking from 130.412 km<sup>2</sup> in 1998 to 85.79 km<sup>2</sup> by 2023, a reduction of 44.62 km<sup>2</sup>. This decline highlights the impact of urban expansion and potential deforestation activities over the years. Water bodies within the study area also experienced a decline, decreasing by 3.099 km<sup>2</sup>, from 9.627 km<sup>2</sup> in 1998 to 6.528 km<sup>2</sup> in 2023. This decline, demonstrating a 0.778% drop in percentage terms, may reflect changes in hydrological conditions, water management practices, or the effects of urbanization. From 1998-2023, bare land increased from 111.774 km<sup>2</sup> to 143.64 km<sup>2</sup>, an increase of 31.866 km<sup>2</sup>. The increase in bare land is noted with dynamic changes in land use patterns and potentially influenced by environmental factors and human activities. This study showed the complex relationship of land cover changes within the area through the dynamic correlation of land use changes. The changes that are seen, particularly in the category of bare land, emphasize how important it is to pay close attention to these changes because they significantly impact the area's overall land composition. These changes highlight the importance of continuously monitoring and analyzing land-use patterns to comprehend the effects of urbanization, environmental changes, and human actions on the landscape. Figures 6.2-6.6 show the complete change scenario (1998-2023) with statistics for each class. Figures 6.2-6.6 show that the built-up area has grown significantly due to the analysis. The built-up would increase from 10 km<sup>2</sup> to 25 km<sup>2</sup> during the study period. The vegetation and water bodies decreased from 130 km<sup>2</sup> to 85 km<sup>2</sup> and 10 km<sup>2</sup> to 6 km<sup>2</sup>, respectively. Meanwhile, the bare land significantly increased from 1998 to 2023 (111 km<sup>2</sup> to 143 km<sup>2</sup>).



Table 6.2 Land Use and Cover Changes in Sialkot, Pakistan (1998-2023).

Class	1998 (km <sup>2</sup> , %)	2003 (km <sup>2</sup> , %)	2008 (km <sup>2</sup> , %)	2013 (km <sup>2</sup> , %)	2018 (km <sup>2</sup> , %)	2023 (km <sup>2</sup> , %)
Vegetation	130.412, 49.79%	127.412, 48.65%	117.412, 44.83%	89.563, 34.19%	123.88, 47.30%	85.79, 32.75%
Built-up Area	10.07, 3.84%	13.07, 4.99%	16.07, 6.13%	20.989, 8.014%	22.21, 8.484%	25.91, 9.89%
Water	9.627, 3.67%	7.627, 2.91%	6.627, 2.53%	7.957, 3.038%	4.488, 1.713%	6.528, 2.49%
Bare Land	111.774, 42.68%	113.774, 43.44%	121.774, 46.49%	143.372, 54.74%	111.28, 42.49%	143.64, 54.85%
<b>Total</b>	<b>261.883, 100%</b>	<b>261.883, 100%</b>	<b>261.883, 100%</b>	<b>261.881, 100%</b>	<b>261.88, 100%</b>	<b>261.88, 100%</b>

### 6.3.1 LULC Change: Every Consecutive Year, 1998-2023

The changes in land use and land cover (LULC) across different periods between 1998 and 2023 are detailed in Table 6.3. For the period 1998-2003, the observed changes were as follows: built-up areas increased by three km<sup>2</sup> (1.145%), vegetation expanded by five km<sup>2</sup> (2.145%), water bodies grew by two km<sup>2</sup> (0.763%), and bare land increased by four km<sup>2</sup> (2.7637%). Between 2003 and 2008, built-up areas and bare land saw further increases of five km<sup>2</sup> (2.145%) and eight km<sup>2</sup> (3.054%), respectively, whereas vegetation and water bodies experienced changes of +7 km<sup>2</sup> (3.818%) and +1 km<sup>2</sup> (0.381%), showing an increase for vegetation contrary to the previous summary and a slight growth for water bodies. From 2008 to 2013, the trend continued with vegetation gaining 27.84 km<sup>2</sup> (10.63%), indicating a significant recovery, and built-up areas increasing by 14.91 km<sup>2</sup> (10.87%). However, water bodies saw a minor reduction of -1.33 km<sup>2</sup> (-0.50%), and bare land expanded by 21.59 km<sup>2</sup> (8.24%). In the subsequent period, 2013-2018, there was a noticeable growth in built-up areas by 16.23 km<sup>2</sup> (10.469%) and a substantial increase in bare land by 32.086 km<sup>2</sup> (12.25%), while vegetation decreased slightly by -34.32 km<sup>2</sup> (-13.10%), and water bodies slightly increased by 3.46 km<sup>2</sup> (1.324%). The most recent period, 2018-2023, showed significant changes, with built-up areas increasing by 23.69 km<sup>2</sup> (15.411%) and bare land by 32.36 km<sup>2</sup> (12.35%). In contrast, vegetation and water bodies experienced declines of -38.09 km<sup>2</sup> (14.547%) and -2.04 km<sup>2</sup> (-0.778%), respectively. Overall, the data illustrates a significant expansion in built-up areas and bare land over the 25 years, with fluctuations in vegetation and water bodies, culminating in a net decrease in vegetation and a slight overall decrease in water bodies by the end of 2023.



Table 6.3 Changes in Land Use and Land Cover Classes from 1998 to 2023 in the Study Area.

Class	Magnitude of change									
	2018-2023		2013-2018		2008-2013		2003-2008		1998-2003	
	Km <sup>2</sup>	%	Km <sup>2</sup>	%	Km <sup>2</sup>	%	Km <sup>2</sup>	%	Km <sup>2</sup>	%
Vegetation	-38.09	14.547	-34.32	-13.10	27.84	10.63	7	3.818	5	2.145
<b>Built-up area</b>	23.69	15.411	16.23	10.469	14.91	10.87	5	2.145	3	1.145
ater	-2.04	-0.778	3.46	1.324	-1.33	-0.50	1	0.381	2	0.763
Bare land	32.36	12.35	32.086	12.25	21.59	8.24	8	3.054	4	2.7637

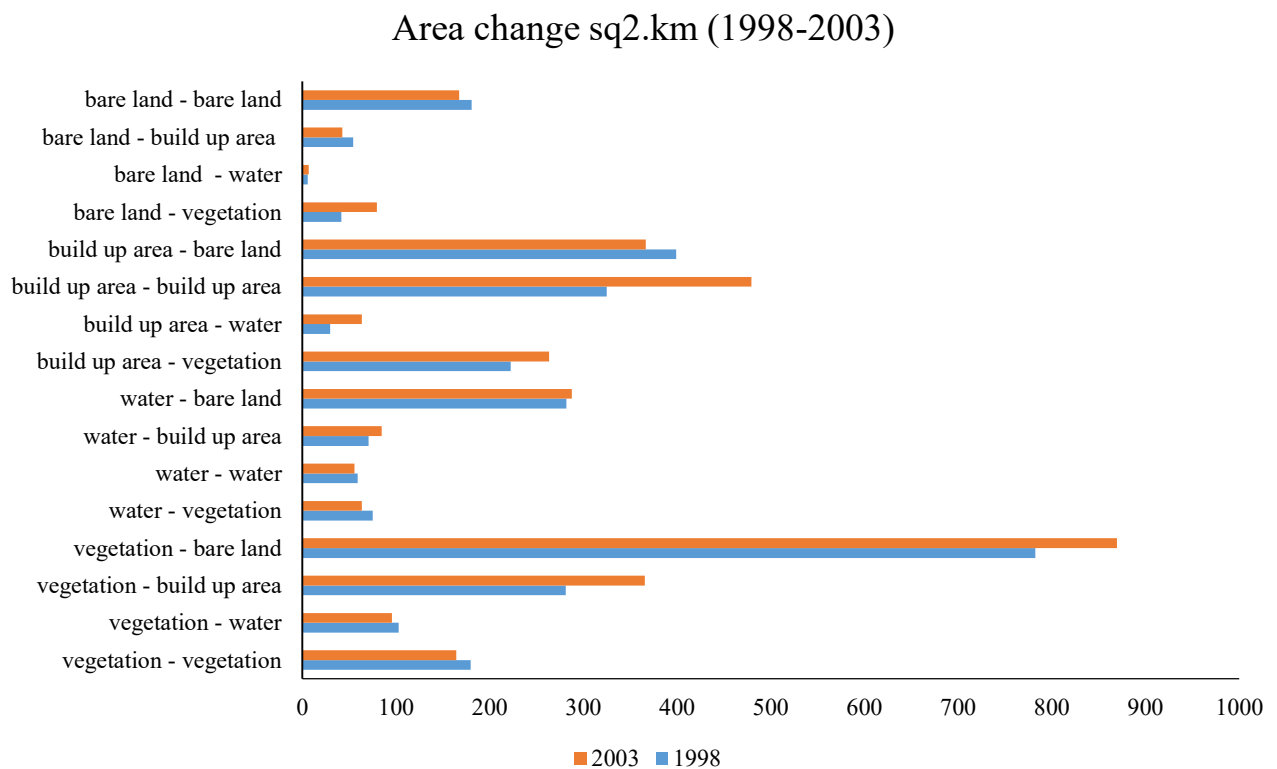


Figure 6.2 Illustrates the collective alterations in Land Use and Land Cover (LULC) areas across Sialkot, as mapped from January 1, 1998, to February 28, 2003, using various Landsat 4 and 5 imagery datasets.



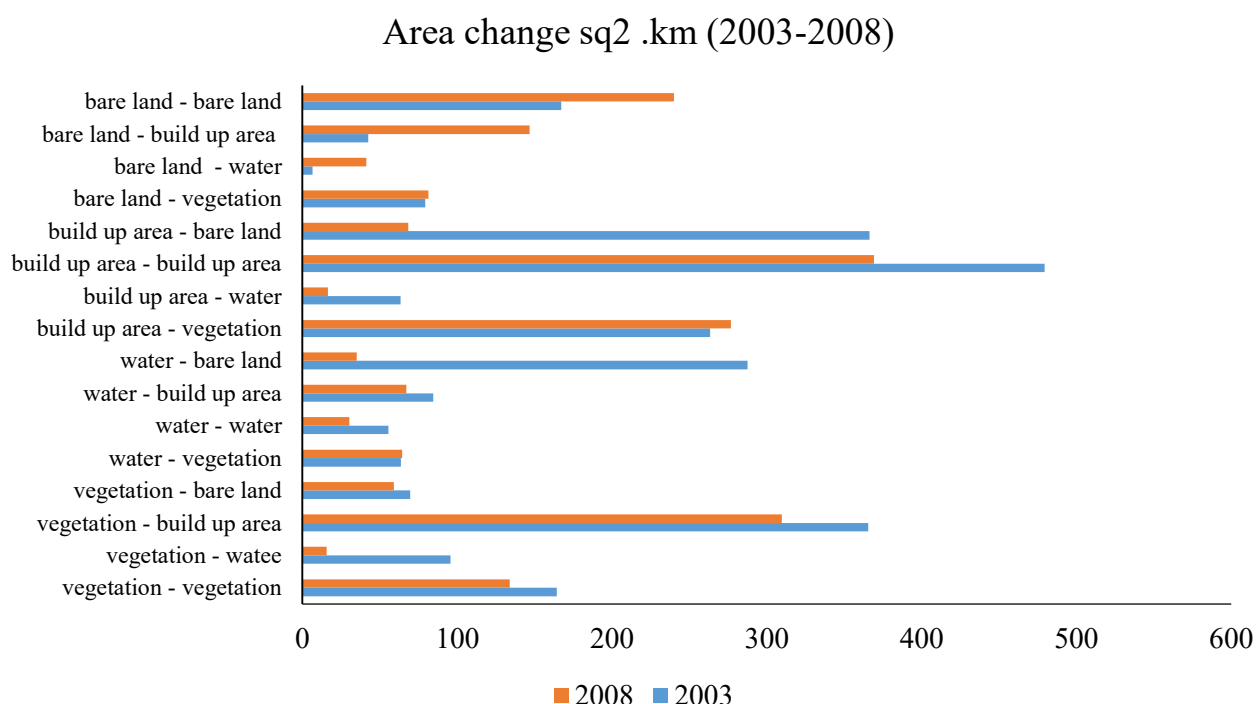


Figure 6.3 Presents the comprehensive transformations in Land Use and Land Cover (LULC) areas from January 1, 2003, to February 28, 2008, utilizing diverse Landsat 4 and 5 imagery datasets within the Sialkot region.

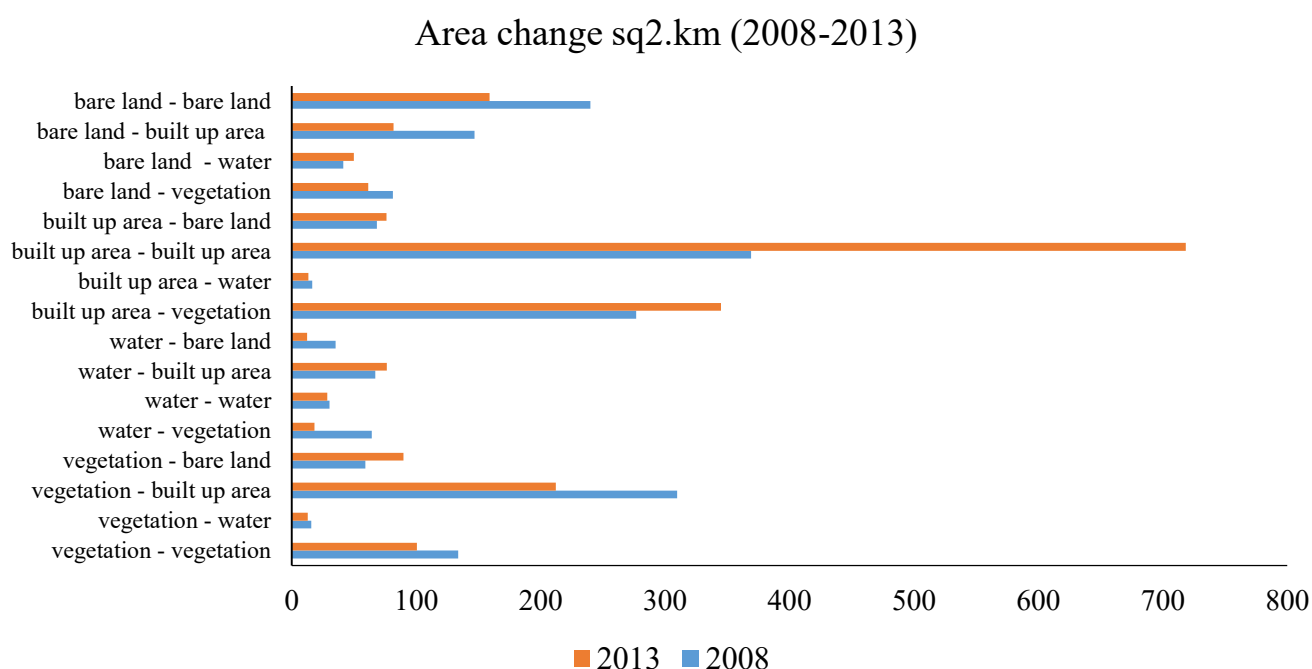


Figure 6. 4 Illustrates the overall changes in Land Use and Land Cover (LULC) areas mapped from January 1, 2008, to February 28, 2013, using various Landsat 4 and 5 imagery datasets in the Sialkot region.



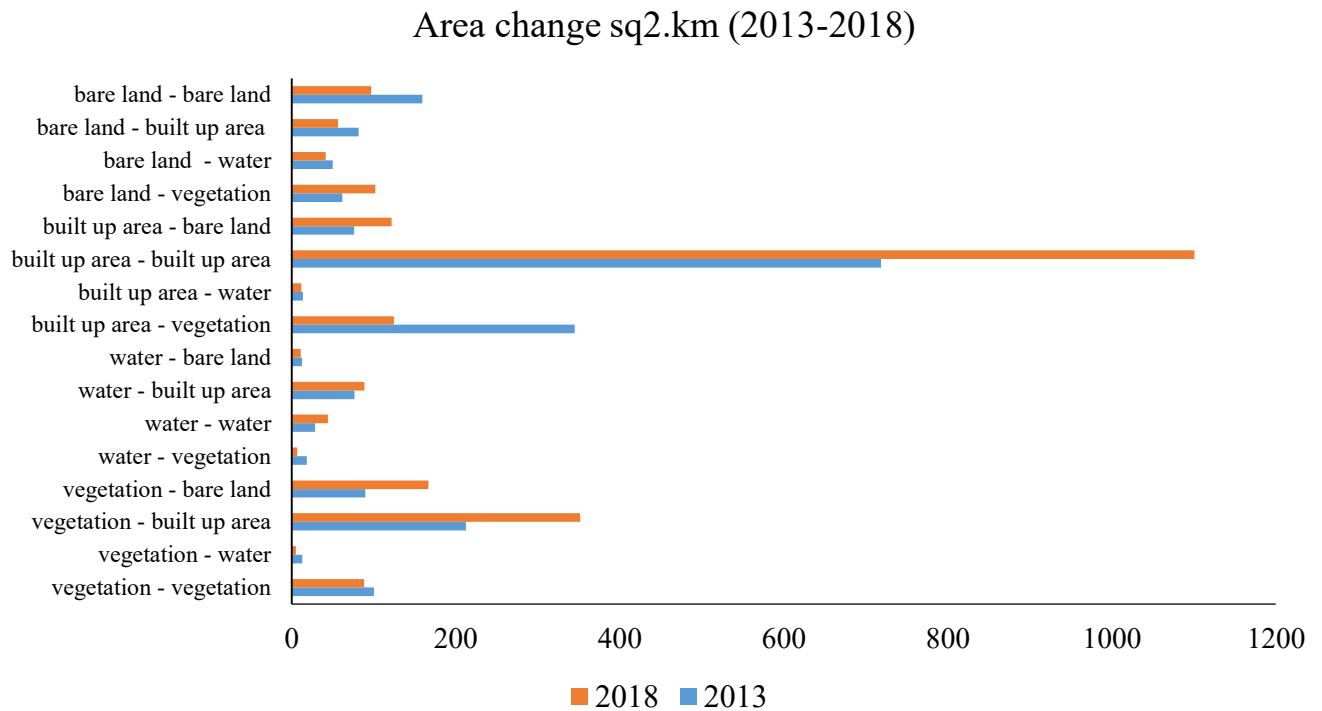


Figure 6.5 Displays the comprehensive alterations in Land Use and Land Cover (LULC) areas mapped from January 1, 2013, to February 28, 2018, utilizing various Landsat 8 imagery datasets within the Sialkot region.

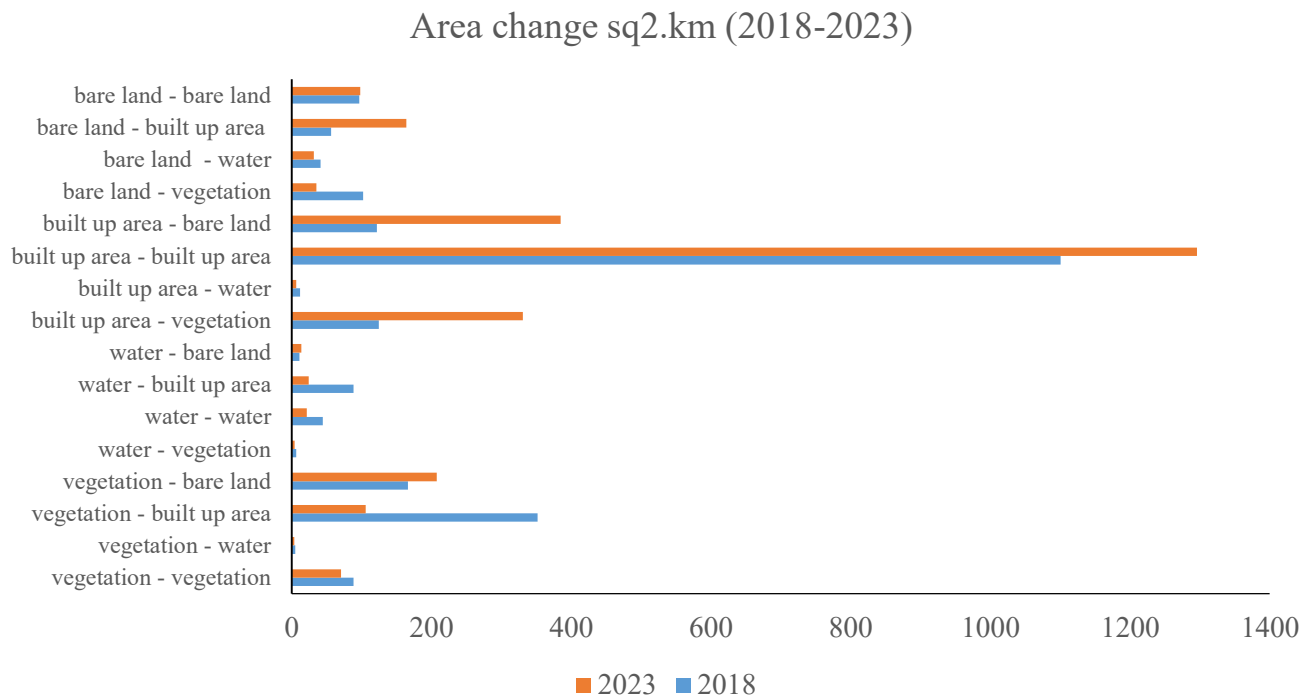


Figure 6. 6 Illustrates the comprehensive changes in Land Use and Land Cover (LULC) areas mapped from January 1, 2018, to February 28, 2023, utilizing various Landsat 8 imagery datasets in the Sialkot region.



Figure 6.2-6.6 showcases the changes in the spatial distribution of Land Use and Land Cover (LULC) categories from 1998 to 2023. Through these visual depictions, the marked rise in developed areas is evident, showcasing the effects of urban and industrial expansion over the twenty-five years. The imagery distinctly shows how Sialkot has undergone significant urban development during this period. Accompanying this, Table 6.3 provides a detailed classification of the satellite imagery used to analyze the study area from 1998 to 2023, further illustrating the transformation within the region.

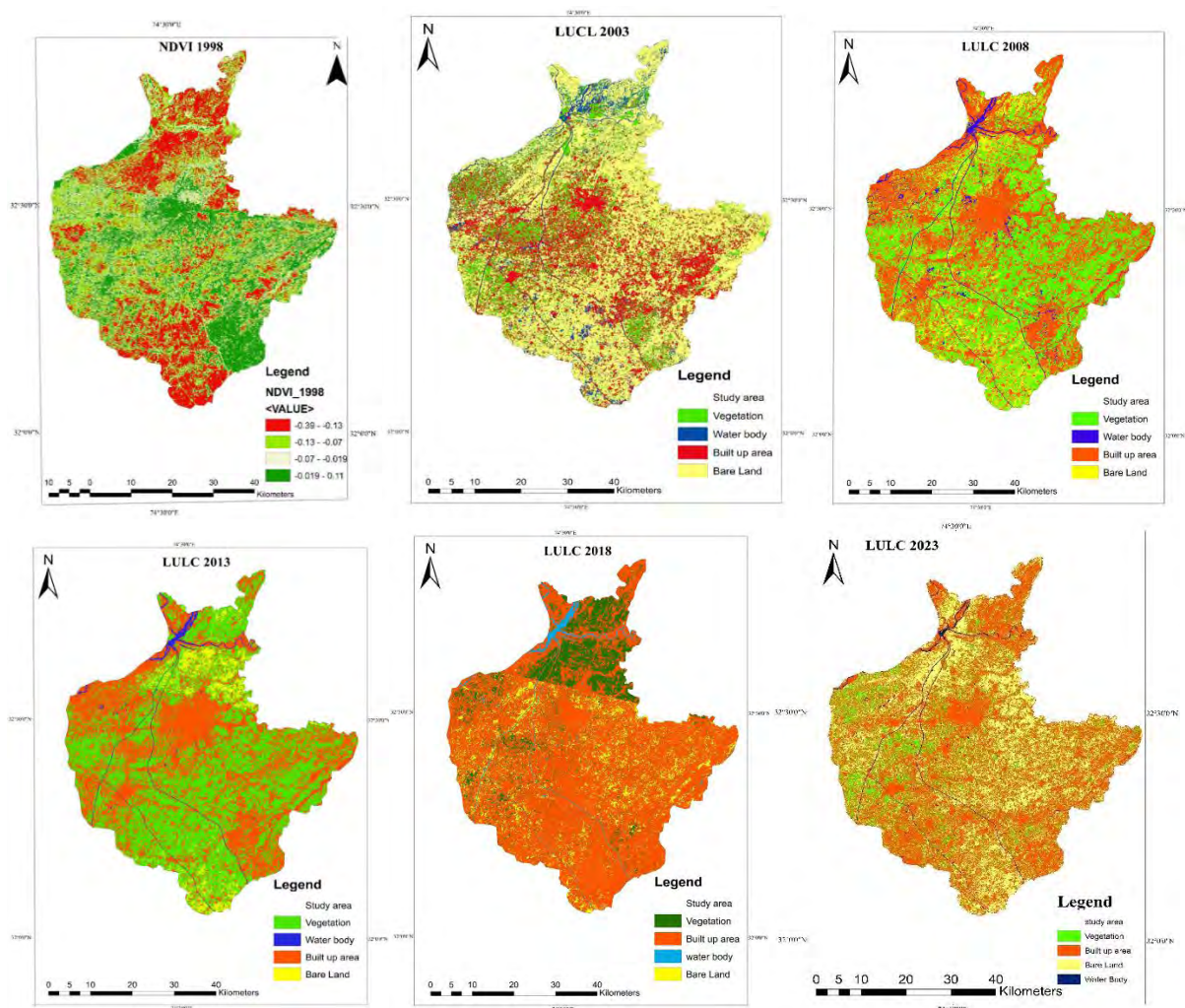


Figure 6.7 A brief study of changes in land use and cover in Sialkot, Pakistan (1998-2023)

### 6.3.2 Accuracy Assessment

The land use and land cover classification in satellite images was conducted using supervised classification techniques focusing on individual pixels. Researchers collected 400 random ground point samples from Landsat maps to assess accuracy, distributing 100 samples per class. These maps resulted from a thorough analysis of Landsat 4, 5, TM, and 8



OLI satellite images taken over 25 years (1998 to 2023) in the Sialkot area, focusing on the vicinity of the Aik stream. This analysis provided valuable insights into how land use and cover changed over time (Table 6.4) due to the availability of datasets from 1998 to 2013 taken Landsat 4 and 5 TM imagery, while from 2013 to 2023 Landsat 8 OLI imagery. This study meticulously documented the classification accuracy for various land cover types, revealing a notable improvement in the detection accuracy of vegetation and built-up areas over time. By 2023, vegetation classification achieved a near-perfect producer accuracy of 99.9%, indicating substantial advancements in remote sensing technologies and methodologies. Built-up areas also exhibited consistently high accuracies, with a significant 95.2% for both producer and user accuracies by 2023, underscoring the model's proficiency in monitoring urban expansion and industrial development. However, classifying water bodies and bare land presented a more complex picture. While water classification reached a user accuracy of 100% in 2008, indicating exceptional model performance in some years, the fluctuating accuracies in these categories reflect the inherent challenges in consistently classifying dynamic and complex landscapes. Land cover changes impacted by seasonal variations and human activities are complicated, as evidenced by the variability of bare land classification accuracy. It appears that Landsat-4, 5, and 8 OLI imagery can be classified fairly accurately using the model's Kappa coefficients, although there were some swings in performance. In 2018 and 2013, the total accuracy peaked, along with the Kappa coefficient, indicating a high level of agreement beyond chance in the classification process, a sign of considerable advancement. This study's findings illustrate the Sialkot region's changing environment, characterized by a significant increase in urbanization and an interplay between natural and anthropogenic influences. Classification accuracy has increased over time due to both the development of satellite imagery analysis and the increasing dependability of these technologies to provide valuable information for sustainable land use, urban planning, and environmental monitoring. This study provides important insight into the accuracy of land cover classification today, laying the foundation for future research and developing policies that protect and manage land resources sustainably.



Table 6.4 The accuracy of Landsat-4, 5, and 8 OLI imagery classification based on the confusion matrix from 1998 to 2023 is presented in this article.

Year	Metric	Vegetation	Built-up Area	Water	Bare Land	Overall Accuracy	Kappa Coefficient
<b>Confusion matrix for the Landsat 4 and 5 TM classification</b>							
<b>1998</b>	Producer Accuracy	88.1%	71.9%	74.7%	50.1%	76.4%	0.659
	User Accuracy	49.0%	91.1%	83.5%	77.2%		
<b>2003</b>	Producer Accuracy	82.8%	77.1%	79.4%	52.3%	74.6%	0.645
	User Accuracy	59.0%	91.1%	89.9%	75.2%		
<b>2008</b>	Producer Accuracy	81.8%	76.1%	78.9%	50.0%	71.6%	0.624
	User Accuracy	64.7%	95.2%	100%	71.4%		
<b>Confusion matrix for Landsat 8 TM classification</b>							
<b>2013</b>	Producer Accuracy	90.0%	95.0%	94.0%	75.0%	89.5%	0.780
	User Accuracy	77.0%	82.0%	76.0%	74.0%		
<b>2018</b>	Producer Accuracy	90.0%	95.0%	94.0%	75.0%	89.5%	0.859
	User Accuracy.	77.0%	82.0%	76.0%	74.0%		
<b>2023</b>	Producer Accuracy.	99.9%	95.2%	78.9%	62.5%	83.5%	0.7801
	User Accuracy	64.7%	95.2%	87.6%	71.4%		

### 6.3.3 NDVI

The vegetation cover of the study area was determined by examining the NDVI values. The NDVI values were produced from the red and infrared bands of Landsat pictures. These data were collected during Sialkot's dry and harvesting seasons, particularly in January and February. The value of NDVI decreased considerably from 1998 to 2023, declining from 0.58 to 0.11. The main reason for the decline was the decrease in green areas and the increase of the built-up area. Comparing NDVI values from 1998 to 2023 suggests that the average NDVI values and the spread of green spaces decreased over time. The results showed that the central part of Sialkot around the Aik Stream of the study area has small NDVI values (Figure 6.8). Therefore, the study established that the vegetation cover has been reduced from



1998 to 2023. From time to time, the decline of vegetation cover contributes to increased surface temperature and industrialization. The (figure 6.9 and Appendix Table 21) is the graph of the NDVI values for each year from 1998 to 2023. The values range between -1 and 1, with higher values meaning higher vegetation health. The maximum NDVI values are 0.79, 0.78, and 0.61 in 1998, 2004, and 2007, corresponding to dense vegetation cover, while the minimum value is 0.10 and 0.11 recorded in multiple years (2016, 2018, and 2023). Annual NDVI values change, demonstrating variation in vegetation density through time.

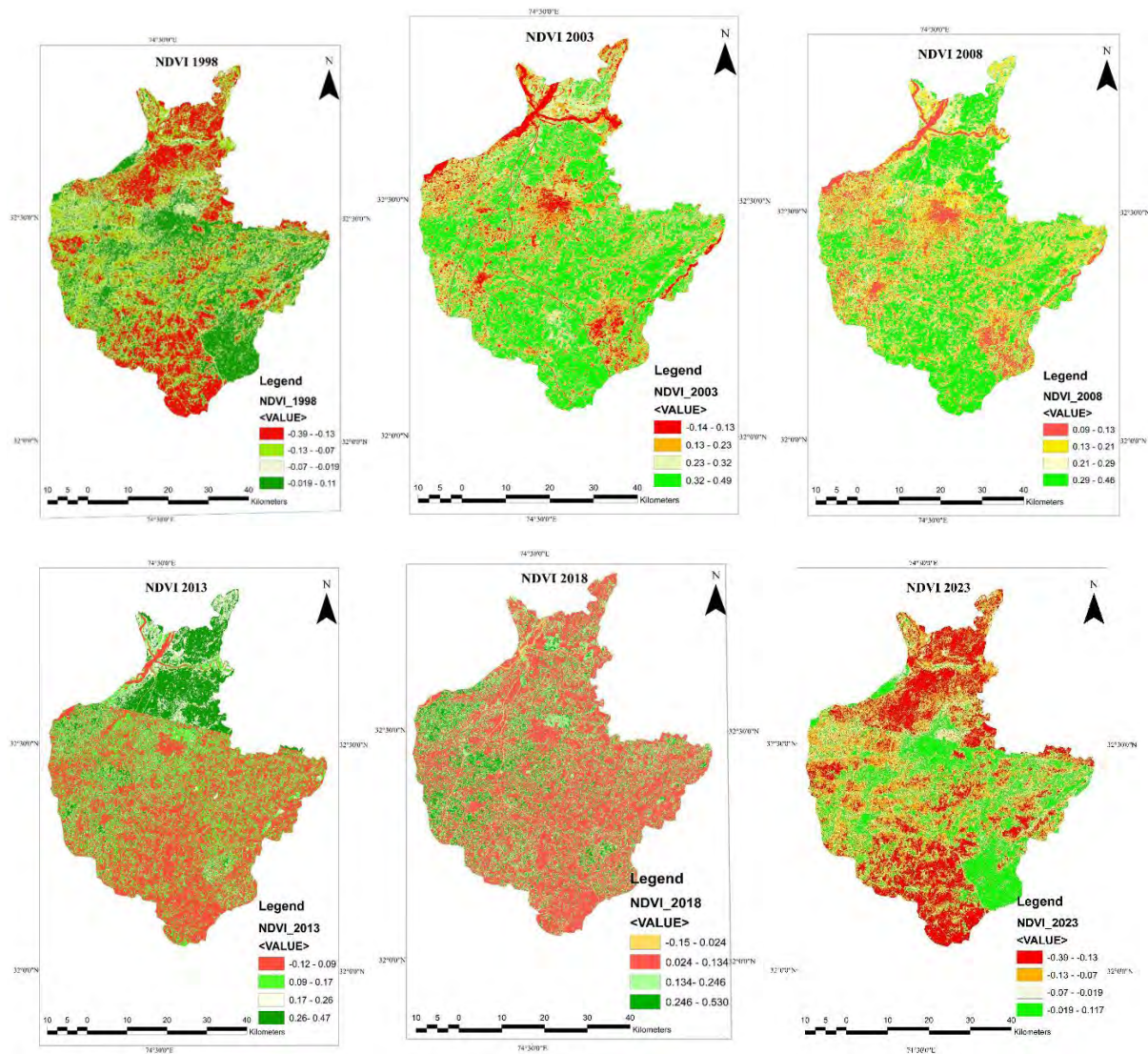


Figure 6.8 Normalized Difference Vegetation Index (NDVI) of the study area from 1998 to 2023



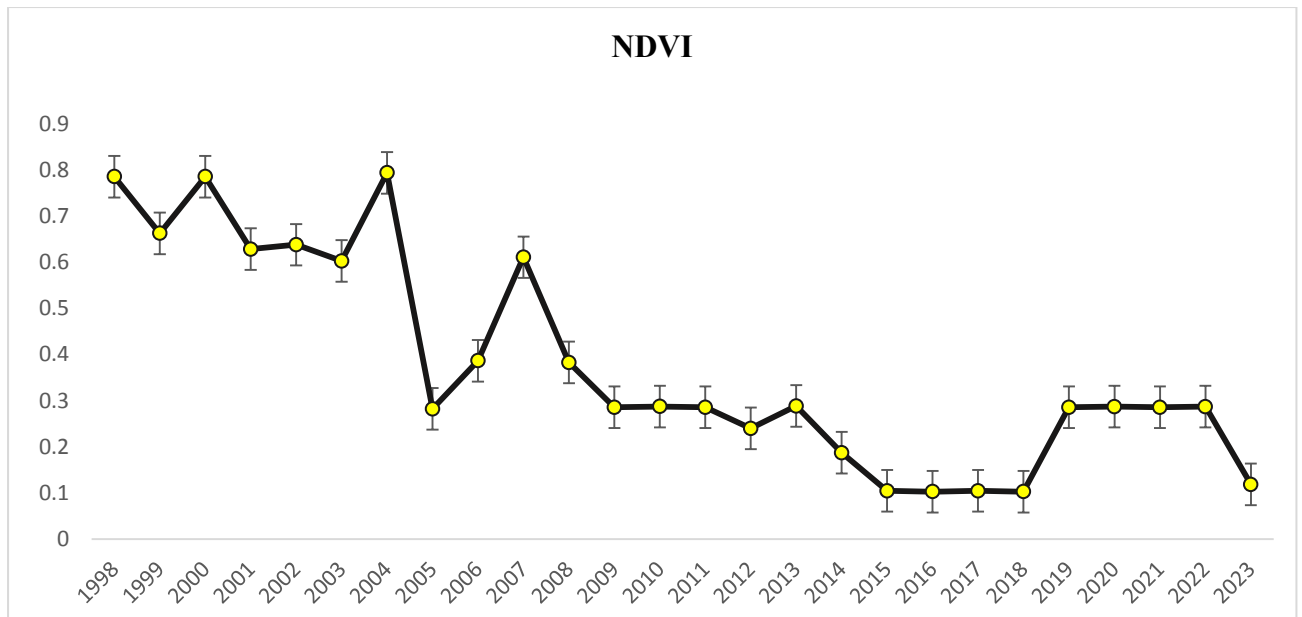


Figure 6. 9 Normalized Difference Vegetation Index (NDVI) of the study area from 1998 to 2023.

#### 6.4 Discussion

This study aimed to assess and illustrate the changes in land use and land cover (LULC) in the City of Sialkot, Pakistan, between 1998 and 2023. For this purpose, satellite images from NASA's Landsat missions were used. A significant increase in urban and industrial development in the city has been observed, as evidenced by an expansion of built-up areas and a trend towards a decline of green spaces that has occurred simultaneously in the city. Population growth has been closely associated with the rapid growth of residential and industrial sectors. Urban landscapes have expanded due to both demographics and economic advancements. Additional insight was provided by the EPA's Built-Up Area Growth Model, which found that the construction of buildings, the expansion of roads, and the development of parking lots are the three leading factors contributing to urbanization, ranked according to their influence (CC, AB, & LDCHN, 2020).

According to the confusion matrix, the model developed in this study was relatively effective in distinguishing built-up (impervious) surfaces from non-built-up (non-impervious) surfaces, albeit with moderate success. This study's confusion matrix illustrates the relative effectiveness of the model. Performance accuracy discrepancies may be explained by using the random forest for the distribution of accuracy assessment points, which resulted in inadvertently favoring non-built-up surfaces (Dong et al., 2020). Therefore, built-up surfaces were likely more susceptible to the impact of classification errors due to the fewer points



designated for them. Within different categories, the error patterns within the model's classification process are distinguished by User Accuracy and Producer Accuracy. The higher the Producer's Accuracy, the more false positives are likely to be generated, whereas the higher the User's Accuracy, the more false negatives are likely to be generated (L. Zhang et al., 2020). Based on the elevated Producer's Accuracy observed on both assessment dates for the classification of built-up surfaces, the model may not be able to distinguish between the spectral responses of varied terrains consistently and accurately, leading to misinterpretation of some non-urban regions as urban built-up spaces (Njoku & Tenenbaum, 2022). The origin of this type of inaccuracy is believed to be mainly linked to regions with exposed dry soils. Using both Landsat 8 OLI and Landsat 5 TM imagery, the zonal statistics accuracy evaluations showed the model was capable of accurately classifying moderately sized urban areas.

NDVI results over 25 years provide an interesting perspective on how vegetation cover in the studied region fluctuates over time. From 1998-2023, the extent of vegetation increased and decreased, reflecting complex interactions between environmental, climatic, and anthropogenic factors. The vegetation cover decreased by 3 km<sup>2</sup> from 1998 to 2003, translating into a modest decrease in percentage terms. As urbanization and industrial development gradually encroached on semi-natural and natural landscapes, this period may reflect their early impacts. As a result of intensified industrial activity and urban sprawl, coupled with possible climatic variations affecting vegetation growth, the subsequent period between 2003 and 2008 saw a more pronounced decline in vegetation cover of 10 km<sup>2</sup>, suggesting an acceleration in land cover change.

This trend was significantly reversed from 2008 to 2013, with vegetation cover rising by 27.84 km<sup>2</sup>. NDVI values have increased due to improved climatic conditions, possibly augmented by conservation efforts or agricultural expansion. As a result, vegetation cover decreased significantly from 2013 to 2018 and 2018 to 2023, respectively, with losses of 34.32 km<sup>2</sup> and 38.09 km<sup>2</sup>. Decreased vegetation cover and health result from prolonged drought conditions, exacerbated by industrial impacts that degrade water quality. Industrial activities have negatively impacted environmental quality, as evidenced by the magnitude of change in vegetation cover. In contaminated water bodies, plant growth and ecosystem health are adversely impacted by pollution, reducing NDVI values (Yengoh, Dent, Olsson, Tengberg, & Tucker III, 2015). Additionally, these findings reflect broader environmental



challenges caused by industrialization, such as habitat destruction and biodiversity loss, which further contribute to the decline in vegetation cover.

Data quality and environmental conditions during data collection for both datasets are acceptable based on accuracy evaluations (Congalton & Green, 2019). An operational land imager (OLI) and Landsat Thematic Mapper (TM) were used to map built-up areas. This data, however, may not provide relevant information about smaller built-up areas (Foody, 2010). As a result of this study, impervious materials dominate or cover most built-up areas. Environmental elements during image capture have also influenced the generation of built-up area datasets. Factors such as water, shadows, and dry soil pose challenges for many classifications approaches due to spectral similarities. Shadows, found in both sets of imagery, result from the sun's angle or the sensor's positioning but were considered in the model without significantly impacting the study's outcomes. The 2013 imagery was affected by a severe drought, leading to more exposed dry soil, reduced river levels, and wildfires.

Similarly, the 1998 imagery showed extensive clear-cut areas, revealing large patches of dry soil. Some dry, compacted soils were misclassified as built-up areas, likely due to their similar spectral signatures. Nonetheless, these instances were not deemed to alter the overall findings significantly.

Noise and errors from the data sources, modeling techniques, and environmental factors likely influenced the results (H. Wu & Li, 2009). The estimated built-up area for both periods is considered conservative despite including dry soil areas, as inferred from the growth in housing and population within the study area. Between 1998 and 2023, there was notable suburban expansion, much of which the model might not have captured according to visual analysis of true-colour imagery. Suburban environments often feature built-up components like roofs, driveways, and sidewalks. Despite efforts to identify built-up areas in suburban regions using true-colour images, accurate detection was hampered by the sensors' limited resolution and the mistaken identification of dry soils as built-up areas.

The path forward in monitoring and understanding the dynamics of vegetation and land use changes necessitates a strategic integration of cutting-edge remote sensing technologies and the refinement of NDVI analysis methodologies (Bellón, Bégué, Lo Seen, De Almeida, & Simões, 2017). Machine learning and artificial intelligence are promising avenues for enhancing the precision of classifying land use and land cover changes and interpreting NDVI maps (Azedou et al., 2023). As a result of technological advancements,



environmental monitoring can be streamlined, revealing nuanced insights into anthropogenic activities and natural habitats (Jafarzadeh, Mahdianpari, Gill, Brisco, & Mohammadimanesh, 2022). The future should be devoted to developing interdisciplinary research initiatives bridging the gaps between environmental science, urban development, climatology, and policy formulation (Wellmann et al., 2020). Creating comprehensive models of urban expansion, climate variability, and agricultural practices can be catalyzed by the synergy between these diverse fields (Frantzeskaki et al., 2019).

It is also crucial to engage policymakers and the wider community to translate insights from remote sensing and NDVI analyses into viable sustainable development strategies (Biehl et al., 2018). In order to realize environmental conservation and sustainable land use planning, it will be essential to communicate complex scientific data to a non-specialist audience (Skidmore, Bijker, Schmidt, & Kumar, 1997). In response to rapid ecological changes, we can respond more proactively to the challenges posed by exploring new technologies, such as drone surveillance and sensor networks. These technologies facilitate continuous, detailed monitoring of ecosystems, enabling immediate response to environmental degradation or natural disasters (Dwivedi, Sreenivas, Ramana, Reddy, & Sankar, 2006).

## **6.5 Conclusions**

This research systematically utilized modern remote sensing and GIS technology to untangle the complicated dynamics of land use and land cover changes in the Sialkot region from 1998 to 2023. The application of LULC classification and the analysis of Landsat 5 and 8 data have helped illuminate the patterns of urban expansion and its encroachment on vital ecological zones. The study's findings indicate that urbanization and industrialization are relentlessly pushing forward across the landscape. There are serious concerns over the potential degradation of riparian buffer capacities and water quality in connection with this expansion, especially within proximity to the Aik-Stream. The study revealed the difficulties associated with accurately mapping suburban sprawl, primarily due to the spectral overlap of non-built-up surfaces, which complicates the identification of urban boundaries. As a result of rapid urban growth, vegetation cover and water bodies have decreased over the study period. The region's physical landscape has been transformed significantly but poses significant threats to its ecological health and sustainability. Built-up areas have increased by 15.84 km<sup>2</sup> while vegetation and water bodies have decreased, indicating a pressing need for informed urban planning and environmental stewardship. Aik-Stream development and



decreased natural buffers necessitate reevaluating current practices to maintain the region's ecological integrity. This study provides valuable insight into the consequences of unchecked urban expansion, emphasizing the need for balanced environmental conservation and development goals. Ecological and hydrological functions of the region depend on sustainable management practices that protect natural habitats, maintain water quality, and ensure their longevity. Such studies must inform policy and planning decisions to ensure a sustainable future for the Sialkot and similar regions worldwide.



## Discussion and Synthesis

Industrial pollution is a major global problem, especially in rapidly developing countries where numerous factories release their effluents directly into nearby water tributaries. Industrial waste products contaminate surface water with highly toxic metals and organic pollutants (Whitehead, Bussi et al. 2018). Global industrial wastewater release rate is projected to double by 2025 (Water 2017; Hutton and Shafahi 2019), an escalation that could pose a major threat to freshwater resources. The Lancet Commission on Pollution and Health identified water pollution as the leading cause of premature deaths worldwide (Landrigan, Fuller et al. 2017). The UN Sustainable Development Goal (SDG) 6.3 aims to improve water quality by minimizing the emission of hazardous chemicals and enhancing recycling and safe reuse by the end of 2030 (Assembly 2015). Pakistan is the 17th most acute case of water scarcity globally, with 79% of its water being deemed unsafe for consumption (Jabeen, Huang et al. 2015). The situation is exacerbated by a mere 1% of industrial wastewater undergoing treatment before being discharged directly into rivers and drains (Fida, Li et al. 2022). Sialkot City in Pakistan, known for its industrial growth, is experiencing a critical water pollution crisis (Khalid, Rizvi et al. 2021). This is primarily attributed to the Aik-Stream, which has been severely impacted by untreated waste, especially from the leather industries (Naeem, Khalid et al. 2021). The Cleaner Production Centre (CPC) in Sialkot estimates that the ideal leather production level is approximately 297 tons daily, resulting in 9,388 cubic meters of tannery effluent daily. Each tannery in the Sialkot district generates between 547 and 814 cubic meters of effluent per day, culminating in 11,000 cubic meters (Qadir, Malik et al. 2008). In contrast, units manufacturing surgical instruments produce effluents containing acids and dissolved metals from electroplating processes, but the volume of these effluents is significantly less than that from tanneries. These surgical instruments and sports goods contribute minimally to effluent production.

Beyond industrial contamination, the stream's ecological health faces further degradation due to municipal sewage and surface runoff. These factors collectively disrupt its physical, chemical, and biological integrity. This environmental challenge extends beyond the stream, impacting the surrounding habitats, including soil and vegetation, necessitating a comprehensive approach to assessing and addressing these interconnected ecological issues.



Spatial and temporal variations provided a better understanding of stream water pollutants dynamics (Alberto, del Pilar et al. 2001). The first objective of the present study was to highlight the spatial and temporal variations in stream water quality and source identification of contaminants. Multivariate statistical techniques such as HACA, ANOVA and PCA segregated important variables that explained most of the spatial and temporal variations and highlighted the contribution of contamination from different sources. Spatial variations in water quality were identified as the results of anthropogenic stress along the longitudinal stretch of the studied streams, while seasonal variations were primarily derived from natural factors such as rainfall. Using HACA, three classes of sampling sites were distinguished based on the water quality parameters of the stream water: Less Polluted Zone (LPZ), Highly Polluted Zone (HPZ), and Moderately Polluted Zone (MPZ). The LPZ, comprising upstream sites, exhibited the least contamination, whereas the HPZ was notably more contaminated. Principal Component Analysis (PCA) revealed that the higher levels of heavy metals in the mid-stream region were largely attributable to tannery effluents, municipal sewage, and waste from surgical instrument manufacturing units. The presence of municipal and industrial effluents escalates levels of Chemical Oxygen Demand (COD), Biological Oxygen Demand (BOD), nitrogenous compounds, and heavy metals. The concentrations of metals such as Chromium (Cr), Lead (Pb), Cadmium (Cd), and Nickel (Ni) in these areas were found to exceed permissible levels. This significant presence of heavy metals in the stream water raised major concerns for the aquatic life and the biota in the vicinity of the studied stream. In addition to conventional techniques, machine learning models were integrated to predict the Water Quality Index (WQI), achieving faster results with a reduced set of variables. These models proved instrumental in enhancing the efficiency of water quality surveillance through improved monitoring, assessment, and prediction capabilities. The adoption of ML models for water quality evaluation gained traction, noted for their proficiency in rapidly and accurately detecting changes in water conditions. The research aimed to advance this study area by analyzing the effectiveness of various ML models in water quality prediction. In this context, two models, Gradient Boosting (GB) and Random Forests (RF), were utilized to predict the WQI of Aik-Stream. The findings indicated high competence in WQI forecasting for both models, with the RF model having a slight edge over the GB model in predictive accuracy. Both models were noted for their reliability in prediction. Notably, the GB model exhibited its highest accuracy with a limited set of seven input variables: Chemical Oxygen Demand (COD), Total Organic Carbon (TOC), Oil and Grease (OG), Ammonia Nitrogen (NH<sub>3</sub>N), Arsenic (Ar), Nickel (Ni), and



Zinc (Zn). This significant observation suggested the GB model's capacity to assist water managers and policymakers calculate the WQI for rivers and streams more efficiently with fewer variables. This improved methodology offered several advantages, such as decreased processing time, lowered expenses, and reduced reliance on extensive, real-time monitoring across numerous contaminated locations. The study's progression in water resource management tactics aided in improving water quality evaluations, fostering more sustainable and proficient practices in river water management.

The second objective of our research was to analyze the bioaccumulation patterns of heavy metals in the soil surrounding the Aik-Stream. A detailed investigation was conducted to understand the distribution of these toxic metals in various areas adjacent to the stream. Through the application of Self-Organizing Map (SOM) analysis, significant variations in the accumulation of different heavy metals across the regions were identified. This analysis proved particularly insightful, revealing a marked contrast in the levels of most heavy metals between the upstream and downstream segments of the stream, although Arsenic (As) and Mercury (Hg) did not exhibit significant spatial variations. The study found that concentrations of heavy metals were predominantly higher in the mid-stream regions compared to the upstream areas of Aik-Stream. This suggested a pattern of increasing accumulation along the course of the stream. Notably, the mid-stream and downstream sites were contaminated with heavy metals, especially Cadmium (Cd), Chromium (Cr), and Lead (Pb), which were associated with industrial effluent and municipal sewage discharged from tanneries and municipalities in the catchment area. To further understand the sources contributing to this accumulation beyond wastewater, Structural Equation Modeling (SEM) was utilized. SEM provided more profound insights into the origins of heavy metal contamination, indicating that industrial effluents and agricultural activities are significant direct sources contributing to elevated levels of heavy metals in the surrounding soil.

Additionally, the analysis revealed that climatic conditions and the inherent characteristics of the soil indirectly influenced the levels of these metals. A comprehensive and nuanced approach to assessing pollution levels was developed by integrating SOM and SEM methodologies. This integrated methodology enabled a more effective and accurate evaluation of the pollution dynamics across different regions surrounding the Aik-Stream, offering a robust framework for understanding the distribution and the contributing factors to heavy metal pollution. This is crucial for implementing targeted environmental protection and remediation strategies. Furthermore, the study applied the Ecological Risk Index to assess the



environmental threat posed by various heavy metals in the soils around the Aik-Stream. The findings highlighted a severe ecological risk, primarily attributed to Cadmium (Cd), Chromium (Cr), and Lead (Pb), which collectively accounted for 87.31% of the total ecological risk index. Cadmium was particularly concerning, exhibiting the highest individual potential ecological risk with a maximum PERI value of 258.5. This underscored the critical need for targeted environmental management strategies focused on mitigating the impact of these specific heavy metals, given their substantial contribution to the overall ecological risk in the area.

The third objective of the study was to describe the plant diversity patterns around the Aik-Stream to understand how pollution impacts the natural flora. The focus was specifically on the flora of ecosystems affected by industrial wastewater pollution. This involved examining the effects of environmental factors on species composition and distribution patterns around the Aik-Stream region in Sialkot, Pakistan. The research identified 182 plant species belonging to 136 genera and 49 families in the areas surrounding Aik-Stream, which varied in pollution levels. These species were categorized into 127 herbs, 26 shrubs, and 29 trees. Notably, the Poaceae family was found to be predominant, followed by families such as Asteraceae, Fabaceae, and Moraceae. The upstream region exhibited more favorable and less disturbed conditions, supporting a rich abundance and diversity of plant life. In contrast, the mid-stream segment, encountering anthropogenic pressures in urban areas, presented hostile conditions for plant life due to significant surface water degradation. As the stream flows through Sialkot city, it receives a substantial volume of industrial effluents, particularly from tanneries and municipal sewage, leading to a marked deterioration in water quality. This degradation adversely affects plant species. The study concluded that municipal and industrial activities exert considerable stress on plant diversity in the mid-stream HPZ compared to other regions. Many plant species cannot withstand such high levels of contaminants and are outcompeted by invasive species that are more tolerant of these harsh conditions. Only species capable of tolerating fluctuations in heavy metals and other pollutants were observed in the downstream sites. Notably, plant species richness increased in the downstream areas as the water quality improved. A comparative improvement in plant diversity was observed in the downstream MPZ of Aik-Stream. The HPZ is characterized by significant water contamination due to point sources such as tanneries, industrial effluents, and municipal sewage from Sialkot City, resulting in highly degraded water quality. In contrast, the MPZ encompasses downstream sites impacted by non-point sources and



domestic sewage from smaller neighboring towns. As the water flows downstream, sedimentation helps reduce pollution levels. Moreover, the confluence of this water with the Palkhu stream at various points downstream aids in improving water quality through natural dilution. This dilution process plays a pivotal role in gradually improving water quality in the region. Despite these improvements, the constraints of pollution stress mean that these sites remain impaired for the re-appearance of native plant species. However, species tolerant of pollution stress may re-appear, indicating resilience in the face of environmental challenges. This study employed a comprehensive suite of multivariate statistical techniques to elucidate the vegetation distribution patterns in the Aik-Stream region. These techniques included Cluster Analysis, Two-way Cluster Analysis, Indicator Species Analysis (ISA), Species Area Curves, Detrended Correspondence Analysis, Canonical Correspondence Analysis, Structural Equation Modeling, and generalized linear modeling. This array of methods facilitated the comparison of multiple classifications and their interrelationships, thereby yielding robust and factual insights from the analyses, as highlighted in previous studies (Khan et al. 2016). ISA, in particular, was instrumental in associating indicator species with specific environmental conditions (Ahmad, Khan et al. 2019). The application of Indicator Species Analysis was pivotal in identifying indicator plant species within the three significant zones surrounding the Aik-Stream. This analysis provided valuable information regarding species' affinity to specific vegetation zones. The calculation of indicator values was based on species abundance data processed in PCORD. For each vegetation layer – tree, shrub, and herb – in each of the identified pollution zones, at least one statistically significant indicator species was selected using ISA. The Monte Carlo Test was then utilized to ascertain the statistical significance of these indicator values, which represented the percentage of perfect indication based on the combined values of relative abundance and frequency (Dufrêne and Legendre 1997). A threshold of 25% indication and a 95% significance level ( $p \leq 0.05$ ) were set as criteria for determining the indicators. These statistical tools and techniques are integral to this research and hold substantial applicability in vegetation ecology. They are particularly effective for investigating the intricate relationships between vegetation dynamics and various factors, including pollution and environmental gradients.

Building on the objectives previously outlined, the present study involved the identification of indicator plant species using Indicator Species Analysis. The fourth objective was to evaluate the phytoremediation potential of these indicator plants within the Highly Polluted Zone (HPZ). This comprehensive method facilitated a comparative analysis of the



phytoremediation capabilities of indicator plants in the HPZ and their physiological responses to soil heavy metal pollution. Bioaccumulation Coefficient (BAC), Translocation Factor (TF), and Biological Concentration Factor (BCF) were employed to measure the standard accumulation, transfer, and concentration quotients for each indicator plant species, as outlined in (Malik, Husain et al. 2010; Malik, Jadoon et al. 2010).

Seventeen species were identified as indicators in the HPZ of Aik-Stream i.e. *Achyranthes aspera*, *Arundo donax*, *Calotropis procera*, *Cannabis sativa*, *Chenopodium album*, *Coronopus didymus*, *Cynodon dactylon*, *Dysphania ambrosioides*, *Eclipta alba*, *Ficus carica*, *Koeleria macrantha*, *Malva neglecta*, *Ricinus communis*, *Parthenium hysterophorus*, *Persicaria glabra*, *Saccharum bengalensis*, and *Ziziphus nummularia*. These species demonstrated significant phytoremediation potential, particularly in managing heavy metal levels. Notably, *Achyranthes aspera* (hyper-accumulator of Ni, Cu), *Calotropis procera* (Pb, Cr), *Cannabis sativa* (Cd, Cu), *Ficus carica* (Cd, Cu), *Ricinus communis* (Pb, Cd), and *Parthenium hysterophorus* (Cd, Cr, Pb) were identified as the top performers out of the 17 indicators due to their maximum capacity for survival in polluted environments. The study further investigated the research question, "How do physiological adaptive responses in indicator plants combine to affect their resilience to metal toxicity in polluted ecosystems?" It explored the correlation between proline accumulation and the health of photosynthetic pigments under heavy metal stress. The findings indicated a significant positive correlation between increased proline levels and tolerance to metal toxicity in the 17 indicator plant species, with a more pronounced effect in the HPZ compared to the LPZ. The study's conclusions highlight the exceptional phytoremediation abilities of several plants, including *Achyranthes aspera*, *Calotropis procera*, *Parthenium hysterophorus*, and *Cannabis sativa*. These species, renowned for their environmental detoxification capacities, are poised to play a vital role in advancing more efficient and diverse phytoremediation strategies. Various studies support this potential. (Linger, Müssig et al. 2002; Usman, Al Jabri et al. 2020; Ejaz, Khan et al. 2022; Sarfraz, Farid et al. 2022; Kaur, Bashir et al. 2023) Utilizing the distinct advantages of each plant could significantly enhance the overall efficacy of phytoremediation methods.

The final objective of this study entailed the deployment of advanced remote sensing and Geographic Information System (GIS) technologies to dissect the complex evolution of land use and land cover (LULC) changes within the Sialkot region, with a concentrated focus on the Aik-Stream vicinity from 1998 to 2023. By strategically employing LULC



classification techniques and analyzing multispectral satellite imagery from the Landsat 5 and 8 missions, we have elucidated the patterns of urban sprawl and its progressive infringement upon critical ecological zones. Our findings bring a considerable surge in built-up areas, highlighting the unyielding force of urbanization and industrial expansion throughout the landscape. This growth, particularly notable in areas close to the Aik-Stream, elicits grave environmental concerns, particularly in relation to the potential compromise of riparian buffer effectiveness and water quality degradation. The research delineates the complexities in accurately mapping the extent of suburban expansion, a task complicated by the spectral overlap between built-up and non-built-up surfaces, which obscures the clear demarcation of urban boundaries.

Moreover, the observed decrease in vegetation cover and water bodies during the study period starkly signifies the environmental repercussions of rapid industrial development. This transformation extends beyond altering the physical landscape, posing substantial threats to the ecological health and sustainability of the region. With an increase of 15.84 km<sup>2</sup> in built-up areas and a significant decline in green spaces and aquatic environments, the study underscores the critical need for well-informed urban planning and robust environmental stewardship. The encroachment of urban development towards the Aik-Stream and the reduction in natural buffers demand a reassessment of existing practices to protect the area's ecological integrity. This research offers invaluable insights into the ramifications of unchecked urban growth, emphasizing the imperative for strategies that reconcile developmental objectives with environmental preservation. It advocates for sustainable management practices aimed at safeguarding natural habitats, ensuring water quality, and maintaining the ecological and hydrological functions of the region.

## **7.1. Conclusion**

This research adopts a comprehensive framework for evaluating the effects of human activities on the natural aquatic ecosystem of Aik-Stream and its vicinity. It offers insights into the consequences of anthropogenic actions on the area's water quality and plant life, serving as a foundational reference for future investigations into the region's ecological facets. The study delves into contaminants' spatial and temporal dynamics and their influence on the ecosystems surrounding Aik-Stream, a previously unexplored subject. The intensification of industrial, urban, and agricultural practices within the Aik-Stream catchment has led to the discharge of effluents laden with toxic pollutants, including heavy



metals, adversely affecting the stream's physical, chemical, and biological integrity. This degradation is most pronounced near point sources where the stream flow is minimal. Vegetation near polluted stream segments is experiencing stress from heavy metal contamination, resulting in a spatial segmentation of plant diversity across upstream, mid-stream, and downstream areas. This disruption facilitates the encroachment of invasive species, significantly altering the ecosystem's structure and function and diminishing biodiversity. Given its longstanding role as a critical water source for human, domestic, and agricultural purposes, the current compromised state of Aik-Stream is untenable. Restoring Aik-Stream is imperative for the local community's welfare and maintaining the ecological and abiotic continuity of the River Chenab. Addressing the deteriorating water quality and the loss of plant diversity near Aik-Stream is urgent, necessitating the restoration of these polluted aquatic ecosystems and safeguarding public health. This critical situation warrants immediate intervention from relevant authorities, including government bodies, municipalities, and environmental agencies, to undertake protective, management, and conservation measures for stream ecosystems. The findings from this study provide a valuable basis for the effective restoration of the Aik-Stream ecosystem, highlighting the pressing need for concerted efforts towards environmental stewardship and sustainable management.

## 7.2 Recommendations and Strategies for Enhancing Stream and Catchment Area

To enhance the ecological health of the Aik-Stream and its adjacent catchment area, the ensuing suggestions are presented for deliberation and implementation:

- ✚ Immediate intervention by relevant authorities, including national, provincial, and district governments and environmental departments, is critical. This includes enacting stringent policies, such as the prohibition of discharging untreated industrial effluents and municipal sewage, to facilitate Aik-Stream rehabilitation.
- ✚ Mitigation of environmental degradation, enforce stringent regulations concerning the generation and disposal of effluents from human activities within the catchment area.
- ✚ Mandate environmental impact assessments before initiating construction projects for industrial facilities, residential developments, and other infrastructural endeavors to preempt ecological disruptions.
- ✚ Implement substantial penalties for organizations, industries, and individuals found to be disposing of waste, irrespective of its state, into the environment as a deterrent against pollution.



- ✚ Introduce a taxation scheme based on the "polluter pays" principle, levying charges on entities and individuals proportional to their environmental impact, to foster a culture of accountability and environmental stewardship.
- ✚ Prioritization of stream management efforts, especially in sections of the stream that traverse urban and industrial zones, to address pollution at its source and restore ecological balance.
- ✚ Launch educational campaigns targeting communities living near Aik-Stream to raise awareness about the detrimental effects of stream pollution on human health and the environment, promoting sustainable practices and community involvement in conservation efforts.
- ✚ These measures, collectively, aim to safeguard the Aik-Stream against further degradation, ensuring its preservation for future generations and maintaining the region's ecological integrity.

## **7.2 For the improvement of stream water quality**

To improve the water quality of Aik-Stream, the following strategic recommendations are proposed:

- ✚ Implementation of comprehensive strategies, including the enactment of robust legislation, rigorous enforcement of existing environmental laws, and proactive stream management practices to mitigate the negative impact of point source pollution on Aik-Stream. Particular attention should be paid to those pollutants that significantly contribute to spatial and temporal water quality variations, ensuring their concentrations remain below internationally recognized environmental standards.
- ✚ Establish and enforce strict guidelines for the on-site treatment of effluents, ensuring that all waste is adequately treated before discharge.
- ✚ Promote the development and operationalization of waste treatment facilities dedicated to processing industrial and municipal waste, enhancing the overall efficacy of waste management practices.
- ✚ Incentivize industrial entities to incorporate effluent treatment solutions at the point of origin. Concurrently, stringent penalties should be imposed on facilities that fail to treat their waste before discharge, ensuring compliance with environmental regulations.



- ✚ Initiate regular monitoring programs targeting the effluent outputs of industrial and municipal sources to assess compliance with treatment standards and environmental impact on stream water quality.
- ✚ Engage community members and stakeholders in educational initiatives focused on raising awareness about the issues of aquatic pollution. These programs should aim to inform the public about the adverse effects of pollution on aquatic ecosystems and the potential risks to human health.
- ✚ Develop comprehensive water quality guidelines applicable not only to Aik-Stream but extendable to other streams and rivers at the national level. This endeavor should involve conducting extensive research to establish local water quality benchmarks that reflect each water body's unique environmental and ecological contexts.

### **7.3 For the protection of natural floral diversity**

The following recommendations are based on the information extracted from results and discussion to improve the conditions for natural floral diversity of the Sialkot region:

- ✚ A regular assessment of metal contents in plant species and crops dwelling around Aik-Stream, and the polluted water of this stream should be used to educate people about the consequences of consuming contaminated crops.
- ✚ Development of management plans to control or eradicate invasive plant species that threaten the native floral diversity of the region, promoting the restoration of native habitats.
- ✚ Establish and maintain vegetated buffer zones along stream banks to filter pollutants, stabilize banks, and provide habitat for wildlife and plant diversity.
- ✚ The stream habitat should be managed to provide better conditions for native flora near the streams.

### **7.4 Future Prospects**

The rapid pace of industrialization, urban development, and agricultural intensification in Pakistan has exerted unprecedented pressure on aquatic ecosystems, highlighting the scarcity of baseline data for these stressed environments. Like the present study, the continuation and expansion of scientific research are vital for guiding environmental conservation and management strategies. Such efforts are crucial for enhancing our understanding of community dynamics in and around aquatic systems.



- ✚ It is vital to conduct extensive surveys to quantify the effluents entering stream systems from diverse point sources. This will enable targeted interventions to mitigate pollution impacts effectively.
- ✚ Identified indicator plants can be used for multipurpose purposes, including reforestation drives and smart habitat plantation.
- ✚ Optimize and enhance the root systems of plants to increase their filtration effectiveness, specifically aimed at capturing and breaking down heavy metals and organic contaminants present in wastewater.
- ✚ Encourage agricultural use of treated wastewater for irrigation, emphasizing methods that safeguard crops from contamination and prevent soil degradation.
- ✚ They are investigating macro-invertebrates, including aquatic species such as fish, crustaceans, and mollusks, as biological indicators provide critical information on environmental shifts over different areas and through changing seasons.
- ✚ Birds and turtles residing near stream environments should be studied further as indicators of the biomagnification of heavy metals and organic pollutants, offering a window into the ecosystem's health.
- ✚ The efficacy and applicability of newly developed evaluation criteria for streams should be tested on other streams within the same region, ensuring their relevance and adaptability.
- ✚ Conducting thorough research on the invasion of exotic species in and around streams is essential. These studies need to evaluate the effects on aquatic ecosystems and the potential risks to native fauna and flora, aiming to grasp the broader ecological consequences.



## References

- Abbaszade, G., Tserendorj, D., Salazar-Yanez, N., Zacháry, D., Völgyesi, P., Tóth, E., & Szabó, C. (2022). Lead and stable lead isotopes as tracers of soil pollution and human health risk assessment in former industrial cities of Hungary. *Applied Geochemistry*, 145, 105397.
- Adejumo, S. A., Awoyemi, V., & Togun, A. O. (2020). Exogenous proline and hormone in combination with compost improves growth and tolerance of maize under heavy metal stress. *Plants Environ*, 2, 40-53.
- Adeoye, A. O., Adebayo, I. A., Afodun, A. M., & Ajijolakewu, K. A. (2022). Benefits and limitations of phytoremediation: Heavy metal remediation review *Phytoremediation* (pp. 227-238): Elsevier.
- Ado, A., Tukur, A. I., Ladan, M., Gumel, S. M., Muhammad, A. A., Habibu, S., & Koki, I. B. (2015). A review on industrial effluents as major sources of water pollution in Nigeria. *Chemistry Journal*, 1(5), 159-164.
- Agarwal, P., Giri, B. S., & Rani, R. (2020). Unravelling the role of rhizospheric plant-microbe synergy in phytoremediation: a genomic perspective. *Current Genomics*, 21(5), 334-342.
- Aggarwal, A., Sharma, I., Tripathi, B., Munjal, A., Baunthiyal, M., Sharma, V. J. P. O. o. r. p., & perspectives, f. (2012). Metal toxicity and photosynthesis. 229-236.
- Ahmad, Z., Khan, S. M., Ali, M. I., Fatima, N., & Ali, S. (2019). Pollution indicandum and marble waste polluted ecosystem; role of selected indicator plants in phytoremediation and determination of pollution zones. *Journal of Cleaner Production*, 236, 117709. doi: <https://doi.org/10.1016/j.jclepro.2019.117709>
- Ahmad, Z., Khan, S. M., Page, S. E., Balzter, H., Ullah, A., Ali, S., . . . Razzaq, A. (2023a). Environmental sustainability and resilience in a polluted ecosystem via phytoremediation of heavy metals and plant physiological adaptations. *Journal of Cleaner Production*, 385, 135733.
- Ahmad, Z., Khan, S. M., Page, S. E., Balzter, H., Ullah, A., Ali, S., . . . Razzaq, A. J. J. o. C. P. (2023b). Environmental sustainability and resilience in a polluted ecosystem via phytoremediation of heavy metals and plant physiological adaptations. 385, 135733.
- Ahmed, M., Mumtaz, R., & Hassan Zaidi, S. M. (2021). Analysis of water quality indices and machine learning techniques for rating water pollution: A case study of Rawal Dam, Pakistan. *Water Supply*, 21(6), 3225-3250.
- Ahriz, S., Nedjraoui, D., & Sadki, N. (2010). The impact of industrial pollution on the ecosystem of Réghaia Lake (Algeria). *Desalination & Water Treatment*, 24.
- Akhtar, F., Li, J., Pei, Y., Xu, Y., Rajput, A., & Wang, Q. (2020). *Optimal features subset selection for large for gestational age classification using gridsearch based recursive feature elimination with cross-validation scheme*. Paper presented at the Frontier Computing: Theory, Technologies and Applications (FC 2019) 8.
- Akhtar, N., Syakir Ishak, M. I., Bhawani, S. A., & Umar, K. (2021). Various natural and anthropogenic factors responsible for water quality degradation: A review. *Water*, 13(19), 2660.
- Akshita, C., Nandini, T., Sumedha, M., & Trishang, U. (2018). Variation in Reducing Sugar and Proline Content of *Saraca asoca* Due to Air Pollution. *Universal Journal of Environmental Research & Technology*, 7(2).
- Al-Yemni, M. N., Sher, H., El-Sheikh, M. A., Eid, E. M. J. S. R., & Essays. (2011). Bioaccumulation of nutrient and heavy metals by *Calotropis procera* and *Citrullus colocynthis* and their potential use as contamination indicators. 6(4), 966-976.



- Alberto, W. D., del Pilar, D. a. M. a., Valeria, A. M. a., Fabiana, P. S., Cecilia, H. A., & de Los Ángeles, B. M. a. (2001). Pattern recognition techniques for the evaluation of spatial and temporal variations in water quality. a case study:: Suqu a River Basin (Córdoba–Argentina). *Water research*, 35(12), 2881-2894.
- Alekseev, I., & Abakumov, E. (2021). Polycyclic aromatic hydrocarbons, mercury and arsenic content in soils of larsemann hills, pravda coast and fulmar Island, Eastern Antarctica. *Bulletin of environmental contamination and toxicology*, 106(2), 278-288.
- Alengebawry, A., Abdelkhalek, S. T., Qureshi, S. R., & Wang, M.-Q. (2021). Heavy metals and pesticides toxicity in agricultural soil and plants: Ecological risks and human health implications. *Toxics*, 9(3), 42.
- Alfeus, A., Molnar, P., Boman, J., Hopke, P. K., & Wichmann, J. (2023). PM2. 5 in Cape Town, South Africa: Chemical characterization and source apportionment using dispersion-normalised positive matrix factorization. *Atmospheric Pollution Research*, 102025.
- Ali, E. A. (1993). Damage to plants due to industrial pollution and their use as bioindicators in Egypt. *Environmental Pollution*, 81(3), 251-255.
- Ali, S., Zhang, S., & Yue, T. (2020). Environmental and economic assessment of rainwater harvesting systems under five climatic conditions of Pakistan. *Journal of Cleaner Production*, 259, 120829.
- Alkhafaji, A., & Nelson, R. A. (2013). *Strategic management: formulation, implementation, and control in a dynamic environment*: Routledge.
- Almeida, C. R. d., Teodoro, A. C., & Gonçalves, A. (2021). Study of the urban heat island (UHI) using remote sensing data/techniques: A systematic review. *Environments*, 8(10), 105.
- Alnaimy, M. A., Shahin, S. A., Vranayova, Z., Zelenakova, M., & Abdel-Hamed, E. M. W. (2021). Long-term impact of wastewater irrigation on soil pollution and degradation: A case study from Egypt. *Water*, 13(16), 2245.
- Alzubi, J., Nayyar, A., & Kumar, A. (2018). *Machine learning from theory to algorithms: an overview*. Paper presented at the Journal of physics: conference series.
- Amabogha, O. N. (2022). *Combining phytoremediation with bioenergy production; exploring options for sustainable remediation*. Middlesex University.
- Amadi, A. N., & Nwankwoala, H. (2013). Evaluation of heavy metal in soils from Enyimba dumpsite in Aba, southeastern Nigeria using contamination factor and geo-accumulation index.
- Anerao, P., Kaware, R., kumar Khedikar, A., Kumar, M., & Singh, L. (2022). Phytoremediation of persistent organic pollutants: Concept challenges and perspectives. *Phytoremediation Technology for the Removal of Heavy Metals and Other Contaminants from Soil and Water*, 375-404.
- Angelakis, A. N., Valipour, M., Ahmed, A. T., Tzanakakis, V., Paranychianakis, N. V., Krasilnikoff, J., . . . Koutsoyiannis, D. (2021). Water conflicts: From ancient to modern times and in the future. *Sustainability*, 13(8), 4237.
- Aparicio, J. D., Garcia-Velasco, N., Urionabarrenetxea, E., Soto, M., Álvarez, A., & Polti, M. A. (2019). Evaluation of the effectiveness of a bioremediation process in experimental soils polluted with chromium and lindane. *Ecotoxicology and environmental safety*, 181, 255-263.
- Appannagari, R. R. (2017). Environmental pollution causes and consequences: a study. *North Asian International Research Journal of Social Science & Humanities*, 3(8), 151-161.
- Arias-Estévez, M., López-Periago, E., Martínez-Carballo, E., Simal-Gándara, J., Mejuto, J.-C., & García-Río, L. (2008). The mobility and degradation of pesticides in soils and



- the pollution of groundwater resources. *Agriculture, ecosystems & environment*, 123(4), 247-260.
- Arif, N., Sharma, N. C., Yadav, V., Ramawat, N., Dubey, N. K., Tripathi, D. K., . . . Sahi, S. J. J. o. P. B. (2019). Understanding heavy metal stress in a rice crop: toxicity, tolerance mechanisms, and amelioration strategies. 62, 239-253.
- Ariffin, M. (2019). Enforcement of environmental pollution control laws: a Malaysian case study. *International Journal of Public Law and Policy*, 6(2), 155-169.
- Arnell, N. W. (1999). Climate change and global water resources. *Global environmental change*, 9, S31-S49.
- Arnon, D. I. (1949). Copper enzymes in isolated chloroplasts. Polyphenoloxidase in *Beta vulgaris*. *Plant physiology*, 24(1), 1.
- Asadollah, S. B. H. S., Sharafati, A., Motta, D., & Yaseen, Z. M. (2021). River water quality index prediction and uncertainty analysis: A comparative study of machine learning models. *Journal of environmental chemical engineering*, 9(1), 104599.
- Asante-Badu, B., Kgorutla, L., Li, S., Danso, P., Xue, Z., & Qiang, G. (2020). PHYTOREMEDIATION OF ORGANIC AND INORGANIC COMPOUNDS IN A NATURAL AND AN AGRICULTURAL ENVIRONMENT: A REVIEW. *Applied Ecology & Environmental Research*, 18(5).
- Aslam, M., Saeed, M. S., Sattar, S., Sajad, S., Sajjad, M., Adnan, M., . . . Sharif, M. T. J. I. J. P. A. B. (2017). Specific role of proline against heavy metals toxicity in plants. 5(6), 27-34.
- Assembly, U. G. (2015). The 2030 agenda for sustainable development. *Resolution: Middlesbrough, UK*.
- Austin, M., Cunningham, R., & Fleming, P. (1984). New approaches to direct gradient analysis using environmental scalars and statistical curve-fitting procedures. *Vegetatio*, 55, 11-27.
- Ayangbenro, A. S., Babalola, O. O. J. I. j. o. e. r., & health, p. (2017). A new strategy for heavy metal polluted environments: a review of microbial biosorbents. 14(1), 94.
- Azedou, A., Amine, A., Kisekka, I., Lahssini, S., Bouziani, Y., & Moukrim, S. (2023). Enhancing Land Cover/Land Use (LCLU) classification through a comparative analysis of hyperparameters optimization approaches for deep neural network (DNN). *Ecological Informatics*, 78, 102333.
- Azizullah, A., Khattak, M. N. K., Richter, P., & Häder, D.-P. (2011). Water pollution in Pakistan and its impact on public health—a review. *Environment international*, 37(2), 479-497.
- Bagherzadeh, F., Mehrani, M.-J., Basirifard, M., & Roostaei, J. (2021). Comparative study on total nitrogen prediction in wastewater treatment plant and effect of various feature selection methods on machine learning algorithms performance. *Journal of Water Process Engineering*, 41, 102033.
- Balali-Mood, M., Naseri, K., Tahergorabi, Z., Khazdair, M. R., & Sadeghi, M. (2021). Toxic mechanisms of five heavy metals: mercury, lead, chromium, cadmium, and arsenic. *Frontiers in pharmacology*, 227.
- Ball, P. (2008). Water as an active constituent in cell biology. *Chemical reviews*, 108(1), 74-108.
- Banerjee, S., Banerjee, A., Palit, D., & Roy, P. (2019). Assessment of vegetation under air pollution stress in urban industrial area for greenbelt development. *International Journal of Environmental Science and Technology*, 16(10), 5857-5870.
- Banerjee, U., & Gupta, S. (2017). Metal contamination in cultivated vegetables and agricultural soils irrigated with untreated industrial wastewater. *Environmental Pollution and Protection*, 2(1), 15-22.



- Bates, L. S., Waldren, R. a., & Teare, I. (1973). Rapid determination of free proline for water-stress studies. *Plant and soil*, 39, 205-207.
- Baxter, J. (2014). Vegetation sampling using the quadrat method. *Lecture: Methods in ECC (BIO221B)*. Available online: <https://www.csus.edu/indiv/b/baxterj/bio%20221b/vegetation%20sampling%20quadrat.pdf> (accessed on 12 March 2019).
- Bayouli, I. T., Bayouli, H. T., Dell'Oca, A., Meers, E., & Sun, J. (2021). Ecological indicators and bioindicator plant species for biomonitoring industrial pollution: Eco-based environmental assessment. *Ecological Indicators*, 125, 107508.
- Behera, B., & Reddy, V. R. (2002). Environment and accountability: Impact of industrial pollution on rural communities. *Economic and Political Weekly*, 257-265.
- Bello, S., Zakari, Y., Ibeanu, I., & Muhammad, B. (2015). Evaluation of heavy metal pollution in soils of Dana steel limited dumpsite, Katsina state Nigeria using pollution load and degree of contamination indices. *American journal of engineering research*, 4(12), 161-169.
- Bellón, B., Bégué, A., Lo Seen, D., De Almeida, C. A., & Simões, M. (2017). A remote sensing approach for regional-scale mapping of agricultural land-use systems based on NDVI time series. *Remote Sensing*, 9(6), 600.
- Berglund, E. Z., Pesantez, J. E., Rasekh, A., Shafiee, M. E., Sela, L., & Haxton, T. (2020). Review of modeling methodologies for managing water distribution security. *Journal of water resources planning and management*, 146(8), 03120001.
- Best, J. (2019). Anthropogenic stresses on the world's big rivers. *Nature Geoscience*, 12(1), 7-21.
- Bhatti, M. T., & Latif, M. (2011). Assessment of water quality of a river using an indexing approach during the low- flow season. *Irrigation and Drainage*, 60(1), 103-114.
- Bhuiyan, T. F., Ahamed, K. U., Nahar, K., Al Mahmud, J., Bhuyan, M. B., Anee, T. I., . . . Biotechnology, A. (2019). Mitigation of PEG-induced drought stress in rapeseed (*Brassica rapa* L.) by exogenous application of osmolytes. 20, 101197.
- Biehl, P. F., Crate, S., Gardezi, M., Hamilton, L., Harlan, S., Hritz, C., & Silva, J. (2018). Innovative tools, methods, and analysis: social science perspectives on climate change, part 3. *Washington, DC: USGCRP Social Science Coordinating Committee*.
- Bindajam, A. A., Mallick, J., Talukdar, S., Islam, A. R. M. T., & Alqadhi, S. (2021). Integration of artificial intelligence-based LULC mapping and prediction for estimating ecosystem services for urban sustainability: past to future perspective. *Arabian Journal of Geosciences*, 14, 1-23.
- Bogardi, J. J., Dudgeon, D., Lawford, R., Flinkerbusch, E., Meyn, A., Pahl-Wostl, C., . . . Vörösmarty, C. (2012). Water security for a planet under pressure: interconnected challenges of a changing world call for sustainable solutions. *Current Opinion in Environmental Sustainability*, 4(1), 35-43.
- Bolan, N. S., Park, J. H., Robinson, B., Naidu, R., & Huh, K. Y. (2011). Phytostabilization: a green approach to contaminant containment. *Advances in agronomy*, 112, 145-204.
- Bondu, R., Cloutier, V., Rosa, E., & Benzaazoua, M. (2016). A review and evaluation of the impacts of climate change on geogenic arsenic in groundwater from fractured bedrock aquifers. *Water, Air, & Soil Pollution*, 227, 1-14.
- Borhannuddin Bhuyan, M., Hasanuzzaman, M., Nahar, K., Mahmud, J. A., Parvin, K., Bhuiyan, T. F., . . . approaches, b. (2019). Plants behavior under soil acidity stress: Insight into morphophysiological, biochemical, and molecular responses. 35-82.
- Bosnic, M., Buljan, J., & Daniels, R. (2000). Pollutants in tannery effluents. *United Nations Industrial Development Organization*, 26.
- Bouzekri, S., El Fadili, H., El Hachimi, M. L., El Mahi, M., & Lotfi, E. M. (2020). Assessment of trace metals contamination in sediment and surface water of quarry



- lakes from the abandoned Pb mine Zaida, High Moulouya-Morocco. *Environment, Development and Sustainability*, 22(7), 7013-7031.
- Brack, W., Dulio, V., Ågerstrand, M., Allan, I., Altenburger, R., Brinkmann, M., . . . Escher, B. I. (2017). Towards the review of the European Union Water Framework Directive: recommendations for more efficient assessment and management of chemical contamination in European surface water resources. *Science of the Total Environment*, 576, 720-737.
- Brandão, M., Martins, F., Accioly, A., Santos, N., Romão, M., & Azevedo, A. (2018). Phytoremediation potential and morphological changes of plants growing in the vicinity of lead smelter plant. *International journal of environmental science and technology*, 15(2), 361-372.
- Brezonik, P. L., King, S. O., & Mach, C. E. (2020). The influence of water chemistry on trace metal bioavailability and toxicity to aquatic organisms *Metal ecotoxicology concepts and applications* (pp. 1-31): CRC Press.
- Bukhari, M. A., Ahmad, Z., Ashraf, M. Y., Afzal, M., Nawaz, F., Nafees, M., . . . Manan, A. (2021). Silicon mitigates drought stress in wheat (*Triticum aestivum* L.) through improving photosynthetic pigments, biochemical and yield characters. *Silicon*, 13, 4757-4772.
- Bunce, R., Bogers, M., Evans, D., Halada, L., Jongman, R., Mucher, C., . . . Olsvig-Whittaker, L. (2013). The significance of habitats as indicators of biodiversity and their links to species. *Ecological indicators*, 33, 19-25.
- Buscaroli, A. (2017). An overview of indexes to evaluate terrestrial plants for phytoremediation purposes. *Ecological Indicators*, 82, 367-380.
- Butt, M. Q., Zeeshan, N., Ashraf, N. M., Akhtar, M. A., Ashraf, H., Afroz, A., . . . Naz, S. (2021). Environmental impact and diversity of protease-producing bacteria in areas of leather tannery effluents of Sialkot, Pakistan. *Environmental Science and Pollution Research*, 28(39), 54842-54851.
- Cachada, A., Rocha-Santos, T., & Duarte, A. C. (2018). Soil and pollution: an introduction to the main issues *Soil pollution* (pp. 1-28): Elsevier.
- Cain, S. A. (1935). Ecological studies of the vegetation of the Great Smoky Mountains. II. The quadrat method applied to sampling spruce and fir forest types. *American Midland Naturalist*, 566-584.
- Cantle, J. E. (1986). *Atomic absorption spectrometry*: Elsevier.
- Cao, J., Xie, C., & Hou, Z. (2022). Ecological evaluation of heavy metal pollution in the soil of Pb-Zn mines. *Ecotoxicology*, 31(2), 259-270.
- Carignan, V., & Villard, M.-A. (2002). Selecting indicator species to monitor ecological integrity: a review. *Environmental monitoring and assessment*, 78, 45-61.
- Carlei, V., & Nuccio, M. (2014). Mapping industrial patterns in spatial agglomeration: A som approach to italian industrial districts. *Pattern Recognition Letters*, 40, 1-10.
- Cataldo, D., & Wildung, R. (1978). Soil and plant factors influencing the accumulation of heavy metals by plants. *Environmental health perspectives*, 27, 149-159.
- CC, A., AB, J., & LDCHN, K. (2020). A GIS-BASED SIMULATION APPLICATION TO MODEL SURFACE RUNOFF LEVEL IN URBAN BLOCKS. *FARU JOURNAL-2020*, 56.
- Chai, W. S., Cheun, J. Y., Kumar, P. S., Mubashir, M., Majeed, Z., Banat, F., . . . Show, P. L. J. J. o. C. P. (2021). A review on conventional and novel materials towards heavy metal adsorption in wastewater treatment application. 296, 126589.
- Chandra, R., Saxena, G., & Kumar, V. (2015). Phytoremediation of environmental pollutants: an eco-sustainable green technology to environmental management. *Advances in biodegradation and bioremediation of industrial waste*, 1-29.



- CHATELAIN, M. (2023). Metal Pollution Influences Habitat Preference in Three Endogeic Earthworms.
- Chaudhry, F. N., & Malik, M. (2017). Factors affecting water pollution: a review. *J. Ecosyst. Ecography*, 7(1), 225-231.
- Chehregani, A., Noori, M., & Yazdi, H. L. (2009). Phytoremediation of heavy-metal-polluted soils: Screening for new accumulator plants in Angouran mine (Iran) and evaluation of removal ability. *Ecotoxicology and environmental safety*, 72(5), 1349-1353.
- Chen, S., Wang, X., & Yao, S. (2023). National water-saving city and its impact on agricultural total factor productivity: A case study of nine provinces along the Yellow River, China. *Journal of Cleaner Production*, 417, 138019.
- Chen, W., Yu, Z., Yang, X., Wang, T., Li, Z., Wen, X., . . . Zhang, C. (2023). Unveiling the role of dissolved organic matter on the Hg phytoavailability in biochar-amended soils. *International Journal of Environmental Research and Public Health*, 20(4), 3761.
- Chernen'kova, T. (2014). Biodiversity of forest vegetation under industrial pollution. *Russian journal of ecology*, 45, 1-10.
- Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, 7, e623.
- China, C. R., Maguta, M. M., Nyandoro, S. S., Hilonga, A., Kanth, S. V., & Njau, K. N. (2020). Alternative tanning technologies and their suitability in curbing environmental pollution from the leather industry: A comprehensive review. *Chemosphere*, 254, 126804.
- Chowdhury, M., Mostafa, M., Biswas, T. K., Mandal, A., & Saha, A. K. (2015). Characterization of the effluents from leather processing industries. *Environmental Processes*, 2, 173-187.
- Chowdhury, M., Mostafa, M., Biswas, T. K., & Saha, A. K. (2013). Treatment of leather industrial effluents by filtration and coagulation processes. *Water Resources and Industry*, 3, 11-22.
- Chu, T.-J., Shih, Y.-J., Shih, C.-H., Wang, J.-Q., Huang, L.-M., & Tsai, S.-C. (2022). Developing a model to select indicator species based on individual species' contributions to biodiversity. *Applied Sciences*, 12(13), 6748.
- Ciobanu, D. (2013). SOIL POLLUTION AND CONTAMINATION. THE LEGAL FRAMEWORK AND ACTUAL TRENDS IN ENVIRONMENTAL POLICIES. *Annals of the University Dunarea de Jos of Galati: Fascicle II, Mathematics, Physics, Theoretical Mechanics*, 36(1).
- Cole, M. M., & Smith, R. F. (1984). Vegetation as indicator of environmental pollution. *Transactions of the Institute of British Geographers*, 477-493.
- Congalton, R. G., & Green, K. (2019). *Assessing the accuracy of remotely sensed data: principles and practices*: CRC press.
- Corona, F., Mulas, M., Baratti, R., & Romagnoli, J. A. (2010). On the topological modeling and analysis of industrial process data using the SOM. *Computers & Chemical Engineering*, 34(12), 2022-2032.
- Correa-Metrio, A., Dechnik, Y., Lozano-García, S., & Caballero, M. (2014). Detrended correspondence analysis: A useful tool to quantify ecological changes from fossil data sets. *Boletín de la Sociedad Geológica Mexicana*, 66(1), 135-143.
- Cousins, S. A., & Lindborg, R. (2004). Assessing changes in plant distribution patterns—indicator species versus plant functional types. *Ecological indicators*, 4(1), 17-27.
- Cox, G. W. (1990). Centripetal and centrifugal incentives in electoral systems. *American Journal of Political Science*, 903-935.



- Culmsee, H., Schmidt, M., Schmiedel, I., Schacherer, A., Meyer, P., & Leuschner, C. (2014). Predicting the distribution of forest habitat types using indicator species to facilitate systematic conservation planning. *Ecological Indicators*, 37, 131-144.
- Cunha, G. K. G., FARIA, B. G., NASCIMENTO, C. W. A., Silva, A. J. D., & CUNHA, K. P. V. (2021). Effects of riparian land use changes on soil attributes and concentrations of potentially toxic elements. *Anais da Academia Brasileira de Ciências*, 93.
- D'Souza, R. J., Varun, M., Masih, J., & Paul, M. S. J. J. o. h. m. (2010). Identification of *Calotropis procera* L. as a potential phytoaccumulator of heavy metals from contaminated soils in Urban North Central India. 184(1-3), 457-464.
- Daily, D. (2006). No step taken to check tanneries pollution (p. 12): March.
- Das, S., & Angadi, D. P. (2020). Land use-land cover (LULC) transformation and its relation with land surface temperature changes: A case study of Barrackpore Subdivision, West Bengal, India. *Remote Sensing Applications: Society and Environment*, 19, 100322.
- Davis III, J. A., & Jacknow, J. (1975). Heavy metals in wastewater in three urban areas. *Journal (Water Pollution Control Federation)*, 2292-2297.
- Dawood, M., Taie, H., Nassar, R., Abdelhamid, M., & Schmidhalter, U. (2014). The changes induced in the physiological, biochemical and anatomical characteristics of *Vicia faba* by the exogenous application of proline under seawater stress. *South African Journal of Botany*, 93, 54-63.
- De Filippis, L., & Pallaghy, C. (1976). The effect of sub-lethal concentrations of mercury and zinc on *Chlorella*: II. Photosynthesis and pigment composition. *Zeitschrift für Pflanzenphysiologie*, 78(4), 314-322.
- De Jong, M., Joss, S., Schraven, D., Zhan, C., & Weijnen, M. (2015). Sustainable-smart-resilient-low carbon-eco-knowledge cities; making sense of a multitude of concepts promoting sustainable urbanization. *Journal of Cleaner production*, 109, 25-38.
- de Vries, W. (2021). Impacts of nitrogen emissions on ecosystems and human health: A mini review. *Current Opinion in Environmental Science & Health*, 21, 100249.
- Del Buono, D., Terzano, R., Panfili, I., & Bartucca, M. L. (2020). Phytoremediation and detoxification of xenobiotics in plants: herbicide-safeners as a tool to improve plant efficiency in the remediation of polluted environments. A mini-review. *International journal of phytoremediation*, 22(8), 789-803.
- Delina, R. E., Arcilla, C., Otake, T., Garcia, J. J., Tan, M., & Ito, A. (2020). Chromium occurrence in a nickel laterite profile and its implications to surrounding surface waters. *Chemical Geology*, 558, 119863.
- Dhanaraj, K., & Angadi, D. P. (2022). Land use land cover mapping and monitoring urban growth using remote sensing and GIS techniques in Mangaluru, India. *GeoJournal*, 87(2), 1133-1159.
- Diarra, I., Kotra, K. K., & Prasad, S. (2021). Assessment of biodegradable chelating agents in the phytoextraction of heavy metals from multi-metal contaminated soil. *Chemosphere*, 273, 128483.
- Dipu, S., Kumar, A. A., & Thanga, S. G. (2012). Effect of chelating agents in phytoremediation of heavy metals. *Remediation Journal*, 22(2), 133-146.
- Dix, R. L. (1961). An application of the point-centered quarter method to the sampling of grassland vegetation. *Rangeland Ecology & Management/Journal of Range Management Archives*, 14(2), 63-69.
- Dixit, R., Wasiullah, X., Malaviya, D., Pandiyan, K., Singh, U. B., Sahu, A., . . . Sharma, P. K. J. S. (2015). Bioremediation of heavy metals from soil and aquatic environment: an overview of principles and criteria of fundamental processes. 7(2), 2189-2212.



- Domene, X., Ramírez, W., Solà, L., Alcañiz, J. M., & Andrés, P. (2009). Soil pollution by nonylphenol and nonylphenol ethoxylates and their effects to plants and invertebrates. *Journal of Soils and Sediments*, 9, 555-567.
- Dong, S., Chen, Z., Gao, B., Guo, H., Sun, D., & Pan, Y. (2020). Stratified even sampling method for accuracy assessment of land use/land cover classification: a case study of Beijing, China. *International Journal of Remote Sensing*, 41(16), 6427-6443.
- Dotaniya, M., Thakur, J., Meena, V., Jajoria, D., & Rathor, G. (2014). Chromium pollution: a threat to environment-a review. *Agricultural Reviews*, 35(2), 153-157.
- Dufrêne, M., & Legendre, P. (1997). Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecological monographs*, 67(3), 345-366.
- Dwivedi, R., Sreenivas, K., Ramana, K., Reddy, P., & Sankar, G. R. (2006). Sustainable development of land and water resources using geographic information system and remote sensing. *Journal of the Indian Society of Remote Sensing*, 34, 351-367.
- Ebrahimi-Khusfi, Z., Nafarzadegan, A. R., & Dargahian, F. (2021). Predicting the number of dusty days around the desert wetlands in southeastern Iran using feature selection and machine learning techniques. *Ecological Indicators*, 125, 107499.
- ED Chaves, M., CA Picoli, M., & D. Sanches, I. (2020). Recent applications of Landsat 8/OLI and Sentinel-2/MSI for land use and land cover mapping: A systematic review. *Remote Sensing*, 12(18), 3062.
- Ejaz, U., Khan, S. M., Aqeel, M., Khalid, N., Sarfraz, W., Naeem, N., . . . Raposo, A. (2022a). Use of Parthenium hysterophorus with synthetic chelator for enhanced uptake of cadmium and lead from contaminated soils—a step toward better public health. *Frontiers in Public Health*, 10, 1009479.
- Ejaz, U., Khan, S. M., Aqeel, M., Khalid, N., Sarfraz, W., Naeem, N., . . . Raposo, A. J. F. i. P. H. (2022b). Use of Parthenium hysterophorus with synthetic chelator for enhanced uptake of cadmium and lead from contaminated soils—a step toward better public health. 10, 1009479.
- Ejaz, U., Khan, S. M., Khalid, N., Ahmad, Z., Jehangir, S., Fatima Rizvi, Z., . . . Raposo, A. (2023a). Detoxifying the heavy metals: a multipronged study of tolerance strategies against heavy metals toxicity in plants. *Frontiers in Plant Science*, 14, 1154571.
- Ejaz, U., Khan, S. M., Khalid, N., Ahmad, Z., Jehangir, S., Fatima Rizvi, Z., . . . Raposo, A. J. F. i. P. S. (2023b). Detoxifying the heavy metals: a multipronged study of tolerance strategies against heavy metals toxicity in plants. 14, 1154571.
- El-Amier, Y. A., Elnaggar, A. A., & El-Alfy, M. A. (2017). Evaluation and mapping spatial distribution of bottom sediment heavy metal contamination in Burullus Lake, Egypt. *Egyptian Journal of Basic and Applied Sciences*, 4(1), 55-66.
- El-Naggar, A., Ahmed, N., Mosa, A., Niazi, N. K., Yousaf, B., Sharma, A., . . . Chang, S. X. (2021). Nickel in soil and water: Sources, biogeochemistry, and remediation using biochar. *Journal of hazardous materials*, 419, 126421.
- Elbana, T., Gaber, H. M., & Kishk, F. M. (2019). Soil chemical pollution and sustainable agriculture. *The soils of Egypt*, 187-200.
- Ertani, A., Mietto, A., Borin, M., & Nardi, S. (2017). Chromium in agricultural soils and crops: a review. *Water, Air, & Soil Pollution*, 228, 1-12.
- Etim, E. J. I. J. E. B. (2012). Phytoremediation and its mechanisms: a review. 2(3), 120-136.
- Fan, Y., Chen, J., Shirkey, G., John, R., Wu, S. R., Park, H., & Shao, C. (2016). Applications of structural equation modeling (SEM) in ecological studies: an updated review. *Ecological Processes*, 5, 1-12.
- Farid, A. M., Lubna, A., Choo, T. G., Rahim, M. C., & Mazlin, M. (2016). A review on the chemical pollution of Langat River, Malaysia. *Asian Journal of Water, Environment and Pollution*, 13(1), 9-15.



- Farooqi, Z. U. R., Hussain, M. M., Ayub, M. A., Qadir, A. A., & Ilic, P. (2022). Potentially toxic elements and phytoremediation: Opportunities and challenges. *Phytoremediation*, 19-36.
- Fazal-ur-Rehman, M. (2019). Polluted water borne diseases: Symptoms, causes, treatment and prevention. *J Med Chem Sci*, 2(1), 21-26.
- Ferreira, M. T., Aguiar, F. C., & Nogueira, C. (2005). Changes in riparian woods over space and time: influence of environment and land use. *Forest Ecology and Management*, 212(1-3), 145-159.
- Fida, M., Li, P., Wang, Y., Alam, S. K., & Nsabimana, A. (2022). Water contamination and human health risks in Pakistan: a review. *Exposure and Health*, 1-21.
- Foody, G. M. (2010). Assessing the accuracy of land cover change with imperfect ground reference data. *Remote Sensing of Environment*, 114(10), 2271-2285.
- Frank, J. J., Poulakos, A. G., Tornero-Velez, R., & Xue, J. (2019). Systematic review and meta-analyses of lead (Pb) concentrations in environmental media (soil, dust, water, food, and air) reported in the United States from 1996 to 2016. *Science of the Total Environment*, 694, 133489.
- Frankel, D. K. (1994). Enforcement of Environmental Laws in Hawai'i. *U. Haw. L. Rev.*, 16, 85.
- Frankel, O. H., Brown, A. H., & Burdon, J. J. (1995). *The conservation of plant biodiversity*: Cambridge University Press.
- Frantzeskaki, N., McPhearson, T., Collier, M. J., Kendal, D., Bulkeley, H., Dumitru, A., . . . Ordóñez, C. (2019). Nature-based solutions for urban climate change adaptation: linking science, policy, and practice communities for evidence-based decision-making. *BioScience*, 69(6), 455-466.
- Freeman, L. A., Corbett, D. R., Fitzgerald, A. M., Lemley, D. A., Quigg, A., & Steppe, C. N. (2019). Impacts of urbanization and development on estuarine ecosystems and water quality. *Estuaries and Coasts*, 42, 1821-1838.
- Gajewska, E., & Skłodowska, M. J. P. G. R. (2008). Differential biochemical responses of wheat shoots and roots to nickel stress: antioxidative reactions and proline accumulation. *54*, 179-188.
- Galiulin, R. V., Bashkin, V. N., Galiulina, R. R., & Birch, P. (2001). A Critical Review: Protection from Pollution by Heavy Metals- Phytoremediation of Industrial Wastewater. *Land Contamination & Reclamation*, 9(4), 349-358.
- Garai, J. (2014). Environmental aspects and health risks of leather tanning industry: a study in the Hazaribag area. *Chinese Journal of Population Resources and Environment*, 12(3), 278-282.
- García-Pardo, K. A., Moreno-Rangel, D., Domínguez-Amarillo, S., & García-Chávez, J. R. (2022). Remote sensing for the assessment of ecosystem services provided by urban vegetation: A review of the methods applied. *Urban Forestry & Urban Greening*, 74, 127636.
- Garelick, H., Jones, H., Dybowska, A., & Valsami-Jones, E. (2008). Arsenic pollution sources. *Reviews of environmental contamination volume 197: International perspectives on arsenic pollution and remediation*, 17-60.
- Garg, S., & Roy, A. (2022). Phytoremediation: An alternative approach for removal of dyes *Phytoremediation* (pp. 369-386): Elsevier.
- Gaur, S., Mittal, A., Bandyopadhyay, A., Holman, I., & Singh, R. (2020). Spatio-temporal analysis of land use and land cover change: a systematic model inter-comparison driven by integrated modelling techniques. *International Journal of Remote Sensing*, 41(23), 9229-9255.



- Gautam, K., Sharma, P., Dwivedi, S., Singh, A., Gaur, V. K., Varjani, S., . . . Ngo, H. H. (2023). A review on control and abatement of soil pollution by heavy metals: Emphasis on artificial intelligence in recovery of contaminated soil. *Environmental Research*, 115592.
- Gavrilescu, M. (2022). Enhancing phytoremediation of soils polluted with heavy metals. *Current Opinion in biotechnology*, 74, 21-31.
- Gazzaz, N. M., Yusoff, M. K., Aris, A. Z., Juahir, H., & Ramli, M. F. (2012). Artificial neural network modeling of the water quality index for Kinta River (Malaysia) using water quality variables as predictors. *Marine pollution bulletin*, 64(11), 2409-2420.
- Geissen, V., Mol, H., Klumpp, E., Umlauf, G., Nadal, M., van der Ploeg, M., . . . Ritsema, C. J. (2015). Emerging pollutants in the environment: a challenge for water resource management. *International soil and water conservation research*, 3(1), 57-65.
- Genchi, G., Carocci, A., Lauria, G., Sinicropi, M. S., & Catalano, A. (2020). Nickel: Human health and environmental toxicology. *International journal of environmental research and public health*, 17(3), 679.
- Genova, G., Della Chiesa, S., Mimmo, T., Borruso, L., Cesco, S., Tasser, E., . . . Niedrist, G. (2022). Copper and zinc as a window to past agricultural land-use. *Journal of Hazardous Materials*, 424, 126631.
- Ghaseminezhad, M., & Karami, A. (2011). A novel self-organizing map (SOM) neural network for discrete groups of data clustering. *Applied soft computing*, 11(4), 3771-3778.
- Ghosh, M., & Singh, S. (2005). A review on phytoremediation of heavy metals and utilization of it's by products. *Asian J Energy Environ*, 6(4), 18.
- Gill, R. A., Kanwar, M. K., Rodrigues dos Reis, A., & Ali, B. (2022). Heavy metal toxicity in plants: recent insights on physiological and molecular aspects. *Frontiers in Plant Science*, 12, 830682.
- Giller, P. S., & Malmqvist, B. (1998). *The biology of streams and rivers*: Oxford University Press.
- Gillispie, E. C., Sowers, T. D., Duckworth, O. W., & Polizzotto, M. L. (2015). Soil pollution due to irrigation with arsenic-contaminated groundwater: Current state of science. *Current Pollution Reports*, 1, 1-12.
- Giuffré, L., Romaniuk, R. I., Marbán, L., Ríos, R. P., & Torres, T. G. (2012). Public health and heavy metals in urban and periurban horticulture. *Emirates Journal of Food and Agriculture*, 148-154.
- Grafkina, M., Pitryuk, A., & Goryacheva, E. (2023). *Soil pollution by heavy metals*. Paper presented at the IOP Conference Series: Earth and Environmental Science.
- Grainger-Brown, J., Malekpour, S., Raven, R., & Taylor, E. (2022). Exploring urban transformation to inform the implementation of the Sustainable Development Goals. *Cities*, 131, 103928.
- Guagliardi, I., Astel, A. M., & Cicchella, D. (2022). Exploring soil pollution patterns using self-organizing maps. *Toxics*, 10(8), 416.
- Günter, S., & Bunke, H. (2003). Validation indices for graph clustering. *Pattern Recognition Letters*, 24(8), 1107-1113.
- Güsewell, S., Peter, M., & Birrer, S. (2012). Altitude modifies species richness–nutrient indicator value relationships in a country-wide survey of grassland vegetation. *Ecological indicators*, 20, 134-142.
- Gworek, B., Dmuchowski, W., & Baczewska-Dąbrowska, A. H. (2020). Mercury in the terrestrial environment: A review. *Environmental Sciences Europe*, 32(1), 1-19.
- Hakanson, L. (1980a). An ecological risk index for aquatic pollution control. A sedimentological approach. *Water research*, 14(8), 975-1001.



- Hakanson, L. (1980b). An ecological risk index for aquatic pollution control: a sedimentological approach. *Water Research*, 14(8), 975-1001. doi: [https://doi.org/10.1016/0043-1354\(80\)90143-8](https://doi.org/10.1016/0043-1354(80)90143-8)
- Hamid, A., Bhat, S. U., & Jehangir, A. (2020). Local determinants influencing stream water quality. *Applied Water Science*, 10(1), 1-16.
- Handique, G., & Handique, A. J. J. E. B. (2009). Proline accumulation in lemongrass (*Cymbopogon flexuosus* Stapf.) due to heavy metal stress. 30(2), 299-302.
- Hasan, S. A., Fariduddin, Q., Ali, B., Hayat, S., & Ahmad, A. (2009). Cadmium: toxicity and tolerance in plants. *J Environ Biol*, 30(2), 165-174.
- Hassan, M. U., Chattha, M. U., Khan, I., Chattha, M. B., Aamer, M., Nawaz, M., . . . Khan, T. A. (2019). Nickel toxicity in plants: reasons, toxic effects, tolerance mechanisms, and remediation possibilities—a review. *Environmental Science and Pollution Research*, 26, 12673-12688.
- Havugimana, E., Bhople, B. S., Kumar, A., Byiringiro, E., Mugabo, J. P., & Kumar, A. (2017). Soil pollution—major sources and types of soil pollutants. *Environmental science and engineering*, 11, 53-86.
- Hayat, M. T., Nauman, M., Nazir, N., Ali, S., & Bangash, N. (2019). Environmental hazards of cadmium: past, present, and future *Cadmium toxicity and tolerance in plants* (pp. 163-183): Elsevier.
- Hayat, S., Hayat, Q., Alyemeni, M. N., Wani, A. S., Pichtel, J., & Ahmad, A. (2012). Role of proline under changing environments: a review. *Plant signaling & behavior*, 7(11), 1456-1466.
- Hazarika, B. B., Gupta, D., Ashu, & Berlin, M. (2020). *A comparative analysis of artificial neural network and support vector regression for river suspended sediment load prediction*. Paper presented at the First International Conference on Sustainable Technologies for Computational Intelligence: Proceedings of ICTSCI 2019.
- Henneberry, S. R., Khan, M. E., & Piewthongngam, K. (2000). An analysis of industrial–agricultural interactions: a case study in Pakistan. *Agricultural Economics*, 22(1), 17-27.
- Heuer, B. J. H. o. p., & stress, c. (2010). Role of proline in plant response to drought and salinity. 3, 213-238.
- Hill, M. O., & Gauch Jr, H. G. (1980). Detrended correspondence analysis: an improved ordination technique. *Vegetatio*, 42(1-3), 47-58.
- Hiscox, J., & Israelstam, G. (1979). A method for the extraction of chlorophyll from leaf tissue without maceration. *Canadian journal of botany*, 57(12), 1332-1334.
- Hoagland, B. W., & Collins, S. L. (1997). Gradient models, gradient analysis, and hierarchical structure in plant communities. *Oikos*, 23-30.
- Hobbs, R. J., & Cramer, V. A. (2008). Restoration ecology: interventionist approaches for restoring and maintaining ecosystem function in the face of rapid environmental change. *Annual Review of Environment and Resources*, 33, 39-61.
- HOCAOĞLU-ÖZYİĞİT, A., & GENÇ, B. N. (2020). Cadmium in plants, humans and the environment. *Frontiers in Life Sciences and Related Technologies*, 1(1), 12-21.
- Hollis, G. (1990). Environmental impacts of development on wetlands in arid and semi-arid lands. *Hydrological Sciences Journal*, 35(4), 411-428.
- Homer, C., Dewitz, J., Yang, L., Jin, S., Danielson, P., Xian, G., . . . Megown, K. (2015). Completion of the 2011 National Land Cover Database for the conterminous United States—representing a decade of land cover change information. *Photogrammetric Engineering & Remote Sensing*, 81(5), 345-354.
- Homer, C. H., Fry, J. A., & Barnes, C. A. (2012). The national land cover database. *US geological survey fact sheet*, 3020(4), 1-4.



- Hoornweg, D., & Pope, K. (2017). Population predictions for the world's largest cities in the 21st century. *Environment and Urbanization*, 29(1), 195-216.
- Hoque, M. N., Tahjib-Ul-Arif, M., Hannan, A., Sultana, N., Akhter, S., Hasanuzzaman, M., . . . Hasan, M. T. (2021). Melatonin modulates plant tolerance to heavy metal stress: Morphological responses to molecular mechanisms. *International Journal of Molecular Sciences*, 22(21), 11445.
- Hou, D., He, J., Lü, C., Ren, L., Fan, Q., Wang, J., & Xie, Z. (2013). Distribution characteristics and potential ecological risk assessment of heavy metals (Cu, Pb, Zn, Cd) in water and sediments from Lake Dalinouer, China. *Ecotoxicology and environmental safety*, 93, 135-144.
- Hu, J., Lin, B., Yuan, M., Lao, Z., Wu, K., Zeng, Y., . . . Zhu, D. (2019). Trace metal pollution and ecological risk assessment in agricultural soil in Dexing Pb/Zn mining area, China. *Environmental geochemistry and health*, 41, 967-980.
- Huang, S., Tang, L., Hupy, J. P., Wang, Y., & Shao, G. (2021). A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing. *Journal of Forestry Research*, 32(1), 1-6.
- Huang, Y., Chen, Q., Deng, M., Japenga, J., Li, T., Yang, X., & He, Z. (2018). Heavy metal pollution and health risk assessment of agricultural soils in a typical peri-urban area in southeast China. *Journal of environmental management*, 207, 159-168.
- Huang, Y., Wang, L., Wang, W., Li, T., He, Z., & Yang, X. (2019). Current status of agricultural soil pollution by heavy metals in China: A meta-analysis. *Science of the Total Environment*, 651, 3034-3042.
- Hull, R. N., Luoma, S. N., Bayne, B. A., Iliff, J., Larkin, D. J., Paschke, M. W., . . . Ward, S. E. (2016). Opportunities and challenges of integrating ecological restoration into assessment and management of contaminated ecosystems. *Integrated environmental assessment and management*, 12(2), 296-305.
- Hussain, M., Khan, S., Abd\_Allah, E., Ul Haq, Z., Alshahrani, T., Alqarawi, A., . . . Ahmad, H. (2019). Assessment of plant communities and identification of indicator species of an Ecotonal Forest zone at Durand Line, District Kurram, Pakistan. *Applied Ecology & Environmental Research*, 17(3).
- Hutton, M. (1983). Sources of cadmium in the environment. *Ecotoxicology and environmental safety*, 7(1), 9-24.
- Hutton, M., & Shafahi, M. (2019). *Water pollution caused by leather industry: a review*. Paper presented at the Energy Sustainability.
- Ilacqua, V., Hänninen, O., Saarela, K., Katsouyanni, K., Künzli, N., & Jantunen, M. (2007). Source apportionment of population representative samples of PM<sub>2.5</sub> in three European cities using structural equation modelling. *Science of the Total Environment*, 384(1-3), 77-92.
- Inglezakis, V., Pouloupoulos, S., Arkhangelsky, E., Zorpas, A., & Menegaki, A. (2016). Aquatic environment *Environment and development* (pp. 137-212): Elsevier.
- Iqbal, M. M., Shoaib, M., Agwanda, P., & Lee, J. L. (2018). Modeling approach for water-quality management to control pollution concentration: A case study of Ravi River, Punjab, Pakistan. *Water*, 10(8), 1068.
- Islam, A. R. M. T., Talukdar, S., Mahato, S., Kundu, S., Eibek, K. U., Pham, Q. B., . . . Linh, N. T. T. (2021). Flood susceptibility modelling using advanced ensemble machine learning models. *Geoscience Frontiers*, 12(3), 101075.
- Islam, M. S., Ahmed, M. K., Raknuzzaman, M., Habibullah-Al-Mamun, M., & Islam, M. K. (2015). Heavy metal pollution in surface water and sediment: a preliminary assessment of an urban river in a developing country. *Ecological indicators*, 48, 282-291.



- Jabeen, A., Huang, X., & Aamir, M. (2015). The challenges of water pollution, threat to public health, flaws of water laws and policies in Pakistan. *Journal of Water Resource and Protection*, 7(17), 1516.
- Jadoon, W. A., & Malik, R. N. (2019). Geochemical approach for heavy metals in suburban agricultural soils of Sialkot, Pakistan. *SN Applied Sciences*, 1(2), 1-11.
- Jafari, J., Khorasani, N., & Danehkar, A. (2010). Ecological risk assessment of lead (Pb) after waste disposal from metallurgical industries. *Research Journal of environmental and earth sciences*, 2(3), 139-145.
- Jafarzadeh, H., Mahdianpari, M., Gill, E. W., Brisco, B., & Mohammadimanesh, F. (2022). Remote Sensing and Machine Learning Tools to Support Wetland Monitoring: A Meta-Analysis of Three Decades of Research. *Remote Sensing*, 14(23), 6104.
- Jahin, H. S., Abuzaid, A. S., & Abdellatif, A. D. (2020). Using multivariate analysis to develop irrigation water quality index for surface water in Kafr El-Sheikh Governorate, Egypt. *Environmental Technology & Innovation*, 17, 100532.
- Jaleel, C. A., Sankar, B., Sridharan, R., & Panneerselvam, R. (2008). Soil salinity alters growth, chlorophyll content, and secondary metabolite accumulation in *Catharanthus roseus*. *Turkish Journal of Biology*, 32(2), 79-83.
- Jan, F. A., Ishaq, M., Ihsanullah, I., & Asim, S. (2010). Multivariate statistical analysis of heavy metals pollution in industrial area and its comparison with relatively less polluted area: a case study from the City of Peshawar and district Dir Lower. *Journal of hazardous materials*, 176(1-3), 609-616.
- Ji, P., Sun, T., Song, Y., Ackland, M. L., & Liu, Y. (2011). Strategies for enhancing the phytoremediation of cadmium-contaminated agricultural soils by *Solanum nigrum* L. *Environmental pollution*, 159(3), 762-768.
- Jiang, H., Guo, H., Sun, Z., Xing, Q., Zhang, H., Ma, Y., & Li, S. (2022). Projections of urban built-up area expansion and urbanization sustainability in China's cities through 2030. *Journal of Cleaner Production*, 367, 133086.
- Jiang, Q., He, J., Ye, G., & Christakos, G. (2018). Heavy metal contamination assessment of surface sediments of the East Zhejiang coastal area during 2012–2015. *Ecotoxicology and environmental safety*, 163, 444-455.
- Jiménez-Valverde, A., Lobo, J., & Hortal, J. (2009). The effect of prevalence and its interaction with sample size on the reliability of species distribution models. *Community Ecology*, 10(2), 196-205.
- Jobbágy, E. G., & Jackson, R. B. (2000). The vertical distribution of soil organic carbon and its relation to climate and vegetation. *Ecological applications*, 10(2), 423-436.
- Johnbull, O., Abbassi, B., & Zytner, R. G. (2019). Risk assessment of heavy metals in soil based on the geographic information system-Kriging technique in Anka, Nigeria. *Environmental Engineering Research*, 24(1), 150-158.
- Johnston, C. A., Bedford, B. L., Bourdaghs, M., Brown, T., Frieswyk, C., Tulbure, M., . . . Zedler, J. B. (2007). Plant species indicators of physical environment in Great Lakes coastal wetlands. *Journal of Great Lakes Research*, 33, 106-124.
- Jones, E. R., van Vliet, M. T. H., Qadir, M., & Bierkens, M. F. P. (2021). Country-level and gridded estimates of wastewater production, collection, treatment and reuse. *Earth Syst. Sci. Data*, 13(2), 237-254. doi: 10.5194/essd-13-237-2021
- Junaid, M., Hashmi, M. Z., & Malik, R. N. (2016). Evaluating levels and health risk of heavy metals in exposed workers from surgical instrument manufacturing industries of Sialkot, Pakistan. *Environmental Science and Pollution Research*, 23, 18010-18026.
- Junaid, M., Hashmi, M. Z., Tang, Y.-M., Malik, R. N., & Pei, D.-S. (2017). Potential health risk of heavy metals in the leather manufacturing industries in Sialkot, Pakistan. *Scientific reports*, 7(1), 8848.



- Kabir, M., Iqbal, M. Z., & Shafiq, M. (2012). Traffic density, climatic conditions and seasonal growth of *Samanea saman* (jacq.) Merr. on different polluted roads of Karachi city. *Pakistan Journal of Botany*, 44(6), 1881-1890.
- Kadi, M. . (2009). "Soil Pollution Hazardous to Environment": A case study on the chemical composition and correlation to automobile traffic of the roadside soil of Jeddah city, Saudi Arabia. *Journal of hazardous materials*, 168(2-3), 1280-1283.
- Kalantar, B., Ueda, N., Idrees, M. O., Janizadeh, S., Ahmadi, K., & Shabani, F. (2020). Forest fire susceptibility prediction based on machine learning models with resampling algorithms on remote sensing data. *Remote Sensing*, 12(22), 3682.
- Kamyab-Talesh, F., Mousavi, S.-F., Khaledian, M., Yousefi-Falakdehi, O., & Norouzi-Masir, M. (2019). Prediction of water quality index by support vector machine: a case study in the Sefidrud Basin, Northern Iran. *Water Resources*, 46, 112-116.
- Kanwar, V. S., Sharma, A., Srivastav, A. L., & Rani, L. (2020). Phytoremediation of toxic metals present in soil and water environment: a critical review. *Environmental Science and Pollution Research*, 27, 44835-44860.
- Kanwar, V. S., Sharma, A., Srivastav, A. L., Rani, L. J. E. S., & Research, P. (2020). Phytoremediation of toxic metals present in soil and water environment: a critical review. 27, 44835-44860.
- Karim, N. (2018). Copper and human health-a review. *Journal of Bahria University Medical and Dental College*, 8(2), 117-122.
- Kastori, R., Plesničar, M., Sakač, Z., Panković, D., & Arsenijević- Maksimović, I. (1998). Effect of excess lead on sunflower growth and photosynthesis. *Journal of Plant Nutrition*, 21(1), 75-85.
- Kathar, S. P., Nagne, A. D., Awate, P. L., & Bhosle, S. (2023). *Comparative Study of Supervised Classification for LULC Using Geospatial Technology*. Paper presented at the International Conference on Soft Computing and its Engineering Applications.
- Kaur, N., Bashir, S., Vadhel, A., Girdhar, M., Malik, T., & Mohan, A. (2023). Assessment of Phytoremediation Potential of Three Weed Plant Species in Soil Contaminated with Lead and Chromium. *International Journal of Agronomy*, 2023.
- Kavitha, D., Gayathri, T., & Devaraj, D. (2023). *Survey on Water Quality Prediction*. Paper presented at the 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS).
- Ke, X., Gui, S., Huang, H., Zhang, H., Wang, C., & Guo, W. (2017). Ecological risk assessment and source identification for heavy metals in surface sediment from the Liaohe River protected area, China. *Chemosphere*, 175, 473-481. doi: <https://doi.org/10.1016/j.chemosphere.2017.02.029>
- Kennedy, J. R., Koth, K. R., & Carruth, R. (2015). Surface and subsurface microgravity data in the vicinity of Sanford Underground Research Facility, Lead, South Dakota: US Geological Survey.
- Khadija, D., Hicham, A., Rida, A., Hicham, E., Nordine, N., & Najlaa, F. (2021). Surface water quality assessment in the semi-arid area by a combination of heavy metal pollution indices and statistical approaches for sustainable management. *Environmental Challenges*, 5, 100230.
- Khalid, N., Rizvi, Z. F., Yousaf, N., Khan, S. M., Noman, A., Aqeel, M., . . . Rafique, A. (2021a). Rising metals concentration in the environment: a response to effluents of leather industries in Sialkot. *Bulletin of Environmental Contamination and Toxicology*, 106(3), 493-500.
- Khalid, N., Rizvi, Z. F., Yousaf, N., Khan, S. M., Noman, A., Aqeel, M., . . . Rafique, A. (2021b). Rising metals concentration in the environment: a response to effluents of



- leather industries in Sialkot. *Bulletin of Environmental Contamination and Toxicology*, 106, 493-500.
- Khan, M. A., & Ghouri, A. M. (2011). Environmental pollution: its effects on life and its remedies. *Researcher World: Journal of Arts, Science & Commerce*, 2(2), 276-285.
- Khan, M. S., Qadir, A., Javed, A., Mahmood, K., Amjad, M. R., & Shehzad, S. (2019). Assessment of aquifer intrinsic vulnerability using GIS based Drastic model in Sialkot area, Pakistan. *International Journal of Economic and Environmental Geology*, 73-84.
- Khan, S. M., Haq, Z. U., Khalid, N., Ahmad, Z., & Ejaz, U. (2022). Utilization of three indigenous plant species as alternative to plastic can reduce pollution and bring sustainability in the environment *Natural Resources Conservation and Advances for Sustainability* (pp. 533-544): Elsevier.
- Khan, S. M., Page, S., Ahmad, H., & Harper, D. (2014). Ethno-ecological importance of plant biodiversity in mountain ecosystems with special emphasis on indicator species of a Himalayan Valley in the northern Pakistan. *Ecological Indicators*, 37, 175-185.
- Khan, W., Khan, S. M., Ahmad, H., Ahmad, Z., & Page, S. (2016). Vegetation mapping and multivariate approach to indicator species of a forest ecosystem: A case study from the Thandiani sub Forests Division (TsFD) in the Western Himalayas. *Ecological Indicators*, 71, 336-351.
- Khatri, N., & Tyagi, S. (2015). Influences of natural and anthropogenic factors on surface and groundwater quality in rural and urban areas. *Frontiers in life science*, 8(1), 23-39.
- Khera, N., Kumar, A., Ram, J., & Tewari, A. (2001). Plant biodiversity assessment in relation to disturbances in mid-elevational forest of Central Himalaya, India. *Tropical Ecology*, 42(1), 83-95.
- Khilchevskiy, V., & Karamushka, V. (2021). Global Water Resources: Distribution and Demand *Clean Water and Sanitation* (pp. 1-11): Springer.
- Khoi, D. N., Quan, N. T., Linh, D. Q., Nhi, P. T. T., & Thuy, N. T. D. (2022). Using machine learning models for predicting the water quality index in the La Buong River, Vietnam. *Water*, 14(10), 1552.
- Khound, N. J., & Bhattacharyya, K. G. (2017). Multivariate statistical evaluation of heavy metals in the surface water sources of Jia Bharali river basin, North Brahmaputra plain, India. *Applied Water Science*, 7, 2577-2586.
- Kiani, S. P., Maury, P., Sarrafi, A., & Grieu, P. (2008). QTL analysis of chlorophyll fluorescence parameters in sunflower (*Helianthus annuus* L.) under well-watered and water-stressed conditions. *Plant science*, 175(4), 565-573.
- Kim, J., de Leeuw, E., Harris-Roxas, B., & Sainsbury, P. (2023). Five urban health research traditions: A meta-narrative review. *Social science & medicine*, 116265.
- Klimek, B., & Niklińska, M. (2007). Zinc and copper toxicity to soil bacteria and fungi from zinc polluted and unpolluted soils: a comparative study with different types of Biolog plates. *Bulletin of environmental contamination and toxicology*, 78, 112-117.
- Kohavi, R., & John, G. H. (1997). Wrappers for feature subset selection. *Artificial intelligence*, 97(1-2), 273-324.
- Kohonen, T. (1990). The self-organizing map. *Proceedings of the IEEE*, 78(9), 1464-1480.
- Kuhn, M., & Johnson, K. (2018). Applied Predictive Modeling, 2nd 2018 Edition: Springer, March.
- Kumar, A., Jigyasu, D. K., Subrahmanyam, G., Mondal, R., Shabnam, A. A., Cabral-Pinto, M., . . . Fagodiya, R. K. (2021). Nickel in terrestrial biota: Comprehensive review on contamination, toxicity, tolerance and its remediation approaches. *Chemosphere*, 275, 129996.



- Kumar, N., Baudh, K., Kumar, S., Dwivedi, N., Singh, D. P., & Barman, S. C. (2013). Accumulation of metals in weed species grown on the soil contaminated with industrial waste and their phytoremediation potential. *Ecological Engineering*, 61, 491-495. doi: <https://doi.org/10.1016/j.ecoleng.2013.10.004>
- Kumar, R. S., & Thambavani, D. S. (2012). Biological monitoring of roadside plants exposed to vehicular pollution in an urban area. *Asian Journal of Research in Chemistry*, 5(10), 1262-1267.
- Kumar, S. S., Kadier, A., Malyan, S. K., Ahmad, A., & Bishnoi, N. R. (2017). Phytoremediation and rhizoremediation: uptake, mobilization and sequestration of heavy metals by plants. *Plant-microbe interactions in agro-ecological perspectives: volume 2: microbial interactions and agro-ecological impacts*, 367-394.
- Kumar, V., Parihar, R. D., Sharma, A., Bakshi, P., Sidhu, G. P. S., Bali, A. S., . . . Gyasi-Agyei, Y. (2019). Global evaluation of heavy metal content in surface water bodies: A meta-analysis using heavy metal pollution indices and multivariate statistical analyses. *Chemosphere*, 236, 124364.
- Kumar, V., Sharma, A., Kaur, P., Sidhu, G. P. S., Bali, A. S., Bhardwaj, R., . . . Cerda, A. (2019). Pollution assessment of heavy metals in soils of India and ecological risk assessment: A state-of-the-art. *Chemosphere*, 216, 449-462.
- Kushwaha, A., Rani, R., Kumar, S., & Gautam, A. J. E. R. (2015). Heavy metal detoxification and tolerance mechanisms in plants: Implications for phytoremediation. 24(1), 39-51.
- Kuussaari, M., Toivonen, M., Heliölä, J., Pöyry, J., Mellado, J., Ekroos, J., . . . Tiainen, J. (2021). Butterfly species' responses to urbanization: differing effects of human population density and built-up area. *Urban Ecosystems*, 24, 515-527.
- LAKEW, D. B. (2020). *REMOTE SENSING BASED URBAN LAND USE/LAND COVER CHANGE DETECTION AND PREDICTION OF ADDIS ABABA CITY*. hu.
- Lal, B. S. (2019). Water for Life: Issues and Challenges. *Int. J. Sci. Res*, 8, 1949-1957.
- Lal, S., Ahmed, N., Srivastava, K., & Singh, D. (2015). Olive (*Olea europaea* L.) seed germination as affected by different scarification treatments. *African Journal of Agricultural Research*, 10(35), 3570-3574.
- Lamb, D. T., Naidu, R., Ming, H., & Megharaj, M. (2012). Copper phytotoxicity in native and agronomical plant species. *Ecotoxicology and environmental safety*, 85, 23-29.
- Landrigan, P., Fuller, R., Acosta, N., Adeyi, O., Arnold, R., & Basu, N. (2017). Baldé, AB; Bertollini, R.; Bose: O'Reilly, S.
- Landrigan, P. J., Fuller, R., Fisher, S., Suk, W. A., Sly, P., Chiles, T. C., & Bose-O'Reilly, S. (2019). Pollution and children's health. *Science of the Total Environment*, 650, 2389-2394.
- Laurent, C., Bravin, M. N., Crouzet, O., Pelosi, C., Tillard, E., Lecomte, P., & Lamy, I. (2020). Increased soil pH and dissolved organic matter after a decade of organic fertilizer application mitigates copper and zinc availability despite contamination. *Science of the Total Environment*, 709, 135927.
- Lavergne, S., Mouquet, N., Thuiller, W., & Ronce, O. (2010). Biodiversity and climate change: integrating evolutionary and ecological responses of species and communities. *Annual review of ecology, evolution, and systematics*, 41, 321-350.
- Lawlor, D. W., & Cornic, G. (2002). Photosynthetic carbon assimilation and associated metabolism in relation to water deficits in higher plants. *Plant, cell & environment*, 25(2), 275-294.
- Lawson, T., Lefebvre, S., Baker, N. R., Morison, J. I., & Raines, C. A. (2008). Reductions in mesophyll and guard cell photosynthesis impact on the control of stomatal responses to light and CO<sub>2</sub>. *Journal of experimental botany*, 59(13), 3609-3619.



- Lee, J., Kaunda, R. B., Sinkala, T., Workman, C. F., Bazilian, M. D., & Clough, G. (2021). Phytoremediation and phytoextraction in Sub-Saharan Africa: Addressing economic and social challenges. *Ecotoxicology and environmental safety*, 226, 112864.
- Leong, W. C., Bahadori, A., Zhang, J., & Ahmad, Z. (2021). Prediction of water quality index (WQI) using support vector machine (SVM) and least square-support vector machine (LS-SVM). *International Journal of River Basin Management*, 19(2), 149-156. doi: 10.1080/15715124.2019.1628030
- Levin, L. A., Boesch, D. F., Covich, A., Dahm, C., Erséus, C., Ewel, K. C., . . . Snelgrove, P. J. E. (2001). The function of marine critical transition zones and the importance of sediment biodiversity. 4, 430-451.
- Li, J., Abdulmohsin, H. A., Hasan, S. S., Kaiming, L., Al-Khateeb, B., Ghareb, M. I., & Mohammed, M. N. (2019). Hybrid soft computing approach for determining water quality indicator: Euphrates River. *Neural Computing and Applications*, 31, 827-837.
- Li, K., Chi, G., Wang, L., Xie, Y., Wang, X., & Fan, Z. (2018). Identifying the critical riparian buffer zone with the strongest linkage between landscape characteristics and surface water quality. *Ecological Indicators*, 93, 741-752.
- Li, X., Jiao, W., Xiao, R., Chen, W., & Chang, A. (2015). Soil pollution and site remediation policies in China: A review. *Environmental Reviews*, 23(3), 263-274.
- Li, Z., Wang, X., Li, J., Zhang, W., Liu, R., Song, Z., . . . Meng, L. (2019). The economic-environmental impacts of China's action plan for soil pollution control. *Sustainability*, 11(8), 2322.
- Liao, Z., Chen, Y., Ma, J., Islam, M. S., Weng, L., & Li, Y. (2019). Cd, Cu, and Zn accumulations caused by long-term fertilization in greenhouse soils and their potential risk assessment. *International journal of environmental research and public health*, 16(15), 2805.
- Licen, S., Astel, A., & Tsakovski, S. (2023). Self-organizing map algorithm for assessing spatial and temporal patterns of pollutants in environmental compartments: A review. *Science of The Total Environment*, 878, 163084.
- Linger, P., Müssig, J., Fischer, H., & Kobert, J. (2002). Industrial hemp (*Cannabis sativa* L.) growing on heavy metal contaminated soil: fibre quality and phytoremediation potential. *Industrial Crops and Products*, 16(1), 33-42.
- Liu, H., Xu, F., Xie, Y., Wang, C., Zhang, A., Li, L., & Xu, H. (2018). Effect of modified coconut shell biochar on availability of heavy metals and biochemical characteristics of soil in multiple heavy metals contaminated soil. *Science of the Total Environment*, 645, 702-709.
- Liu, J., Li, Y., Zhang, B., Cao, J., Cao, Z., & Domagalski, J. (2009). Ecological risk of heavy metals in sediments of the Luan River source water. *Ecotoxicology*, 18, 748-758.
- Lokhande, R. S., Singare, P. U., & Pimple, D. S. (2011). Toxicity study of heavy metals pollutants in waste water effluent samples collected from Talaja industrial estate of Mumbai, India. *Resources and Environment*, 1(1), 13-19.
- Lum, A. F., Ngwa, E., Chikoye, D., & Suh, C. J. I. J. o. P. (2014). Phytoremediation potential of weeds in heavy metal contaminated soils of the Bassa Industrial Zone of Douala, Cameroon. 16(3), 302-319.
- Lumb, A., Sharma, T., & Bibeault, J.-F. (2011). A review of genesis and evolution of water quality index (WQI) and some future directions. *Water Quality, Exposure and Health*, 3, 11-24.
- Madhav, S., Ahamad, A., Singh, A. K., Kushawaha, J., Chauhan, J. S., Sharma, S., & Singh, P. (2020). Water pollutants: sources and impact on the environment and human health. *Sensors in water pollutants monitoring: Role of material*, 43-62.



- Madkour, S. A., & Laurence, J. (2002). Egyptian plant species as new ozone indicators. *Environmental Pollution*, 120(2), 339-353.
- Madlener, R., & Sunak, Y. (2011). Impacts of urbanization on urban structures and energy demand: What can we learn for urban energy planning and urbanization management? *Sustainable Cities and Society*, 1(1), 45-53.
- Mahmood, A., Syed, J. H., Malik, R. N., Zheng, Q., Cheng, Z., Li, J., & Zhang, G. (2014). Polychlorinated biphenyls (PCBs) in air, soil, and cereal crops along the two tributaries of River Chenab, Pakistan: concentrations, distribution, and screening level risk assessment. *Science of the Total Environment*, 481, 596-604.
- Majeed, M., Tariq, A., Haq, S. M., Waheed, M., Anwar, M. M., Li, Q., . . . Jamil, A. (2022). A detailed ecological exploration of the distribution patterns of wild Poaceae from the Jhelum District (Punjab), Pakistan. *Sustainability*, 14(7), 3786.
- Malaeb, Z. A., Summers, J. K., & Pugsek, B. H. (2000). Using structural equation modeling to investigate relationships among ecological variables. *Environmental and Ecological Statistics*, 7, 93-111.
- Malik, R. N., Husain, S. Z., & Nazir, I. (2010). Heavy metal contamination and accumulation in soil and wild plant species from industrial area of Islamabad, Pakistan. *Pak J Bot*, 42(1), 291-301.
- Malik, R. N., Jadoon, W. A., & Husain, S. Z. (2010). Metal contamination of surface soils of industrial city Sialkot, Pakistan: a multivariate and GIS approach. *Environmental geochemistry and health*, 32, 179-191.
- Malone, B. P., Minasny, B., & McBratney, A. B. (2017). *Using R for digital soil mapping* (Vol. 35): Springer.
- Manan, F., Khan, S. M., Muhammad, Z., Ahmad, Z., Abdullah, A., Han, H., . . . Raposo, A. (2022). Floristic composition, biological spectrum, and phytogeographic distribution of the Bin Dara Dir, in the western boundary of Pakistan. *Frontiers in Forests and Global Change*, 5, 1019139.
- Maqbool, A., Ali, S., Rizwan, M., Ishaque, W., Rasool, N., Rehman, M. Z. U., . . . Wu, L. (2018). Management of tannery wastewater for improving growth attributes and reducing chromium uptake in spinach through citric acid application. *Environmental Science and Pollution Research*, 25, 10848-10856.
- Martin, D. C. (1988). *Aquatic habitat of the Tiekel River, southcentral Alaska, and its utilization by resident Dolly Varden (Salvelinus malma)*.
- Martin, J. L., Schottman, R. W., & McCutcheon, S. C. (2018). *Hydrodynamics and transport for water quality modeling*: CRC press.
- Martínez- Alcalá, I., & Bernal, M. P. (2020). Environmental impact of metals, metalloids, and their toxicity. *Metalloids in Plants: Advances and Future Prospects*, 451-488.
- Mateos, B., Conrad-Billroth, C., Schiavina, M., Beier, A., Kontaxis, G., Konrat, R., . . . Pierattelli, R. (2020). The ambivalent role of proline residues in an intrinsically disordered protein: from disorder promoters to compaction facilitators. *Journal of molecular biology*, 432(9), 3093-3111.
- Máthé-Gáspár, G., & Anton, A. (2005). Phytoremediation study: Factors influencing heavy metal uptake of plants. *Acta Biologica Szegediensis*, 49(1-2), 69-70.
- McCutcheon, S., & Schnoor, J. (2003). Overview of phytotransformation and control of wastes. *Phytoremediation: Transformation and control of contaminants*, 1-58.
- McLaughlin, M. J., & Singh, B. R. (1999). *Cadmium in soils and plants: a global perspective*: Springer.
- Mehdizadeh, S., Fathian, F., Safari, M. J. S., & Khosravi, A. (2020). Developing novel hybrid models for estimation of daily soil temperature at various depths. *Soil and Tillage Research*, 197, 104513.



- Mele, P. M., & Crowley, D. E. (2008). Application of self-organizing maps for assessing soil biological quality. *Agriculture, Ecosystems & Environment*, 126(3-4), 139-152.
- Mienye, I. D., & Sun, Y. (2022). A survey of ensemble learning: Concepts, algorithms, applications, and prospects. *IEEE Access*, 10, 99129-99149.
- Minari, G. D., Saran, L. M., Constancio, M. T. L., da Silva, R. C., Rosalen, D. L., de Melo, W. J., & Alves, L. M. C. (2020). Bioremediation potential of new cadmium, chromium, and nickel-resistant bacteria isolated from tropical agricultural soil. *Ecotoxicology and Environmental Safety*, 204, 111038.
- Mishra, B. K., Kumar, P., Saraswat, C., Chakraborty, S., & Gautam, A. (2021). Water security in a changing environment: Concept, challenges and solutions. *Water*, 13(4), 490.
- Mishra, R. K. (2023). Fresh water availability and its global challenge. *British Journal of Multidisciplinary and Advanced Studies*, 4(3), 1-78.
- Mishra, R. K., Mohammad, N., & Roychoudhury, N. (2016). Soil pollution: Causes, effects and control. *Van Sangyan*, 3(1), 1-14.
- Miskowiec, P., Laptas, A., & Zieba, K. (2015). Soil pollution with heavy metals in industrial and agricultural areas: a case study of Olkusz District. *Journal of Elementology*, 20(2).
- Mohamed, M., & Plante, R. (2002). *Remote sensing and geographic information systems (GIS) for developing countries*. Paper presented at the IEEE International Geoscience and Remote Sensing Symposium.
- Mokarram, M., Saber, A., & Sheykhi, V. (2020). Effects of heavy metal contamination on river water quality due to release of industrial effluents. *Journal of Cleaner Production*, 277, 123380.
- Mokhtar, A., Jalali, M., He, H., Al-Ansari, N., Elbeltagi, A., Alsafadi, K., . . . Rodrigo-Comino, J. (2021). Estimation of SPEI meteorological drought using machine learning algorithms. *IEEE Access*, 9, 65503-65523.
- Moriarty, M. M., Koch, I., Gordon, R. A., & Reimer, K. J. (2009). Arsenic speciation of terrestrial invertebrates. *Environmental Science & Technology*, 43(13), 4818-4823.
- Morin, G., & Calas, G. (2006). Arsenic in soils, mine tailings, and former industrial sites. *Elements*, 2(2), 97-101.
- Muangthong, S., & Shrestha, S. (2015). Assessment of surface water quality using multivariate statistical techniques: case study of the Nampong River and Songkhram River, Thailand. *Environmental monitoring and assessment*, 187, 1-12.
- Mulenga, C., Clarke, C., & Meincken, M. (2020). Physiological and growth responses to pollutant-induced biochemical changes in plants: A review.
- Mulk, S., Korai, A. L., Azizullah, A., & Khattak, M. N. K. (2016). Decreased fish diversity found near marble industry effluents in River Barandu, Pakistan. *Ecotoxicology*, 25, 132-140.
- Murayama, Y., Simwanda, M., & Ranagalage, M. (2021). Spatiotemporal analysis of urbanization using GIS and remote sensing in developing countries (Vol. 13, pp. 3681): MDPI.
- Muriithi, F. K. (2016). Land use and land cover (LULC) changes in semi-arid sub-watersheds of Laikipia and Athi River basins, Kenya, as influenced by expanding intensive commercial horticulture. *Remote Sensing Applications: Society and Environment*, 3, 73-88.
- Murtazaa, G., & Usmanb, M. (2022). Assessment of various heavy metals level in groundwater and soil at tannery manufacturing areas of three mega cities (Sialkot, Lahore and Karachi) of Pakistan. *DESALINATION AND WATER TREATMENT*, 266, 121-130.



- Mustapha, A., Aris, A. Z., Yusoff, F. M., Zakaria, M. P., Ramli, M. F., Abdullah, A. M., . . . Narany, T. S. (2014). Statistical approach in determining the spatial changes of surface water quality at the upper course of Kano River, Nigeria. *Water Quality, Exposure and Health*, 6, 127-142.
- Muthusaravanan, S., Sivarajasekar, N., Vivek, J., Paramasivan, T., Naushad, M., Prakashmaran, J., . . . Al-Duaij, O. K. J. E. c. l. (2018). Phytoremediation of heavy metals: mechanisms, methods and enhancements. *16*, 1339-1359.
- Muthusaravanan, S., Sivarajasekar, N., Vivek, J., Vasudha Priyadharshini, S., Paramasivan, T., Dhakal, N., & Naushad, M. (2020). Research updates on heavy metal phytoremediation: enhancements, efficient post-harvesting strategies and economic opportunities. *Green materials for wastewater treatment*, 191-222.
- Naeem, N., Khalid, N., Sarfraz, W., Ejaz, U., Yousaf, A., Rizvi, Z. F., . . . Toxicology. (2021). Assessment of lead and cadmium pollution in soil and wild plants at different functional areas of Sialkot. *107*(2), 336-342.
- Nakagawa, K., Yu, Z.-Q., Berndtsson, R., & Hosono, T. (2020). Temporal characteristics of groundwater chemistry affected by the 2016 Kumamoto earthquake using self-organizing maps. *Journal of Hydrology*, 582, 124519.
- Nakamura, A., Burwell, C. J., Ashton, L. A., Laidlaw, M. J., Katabuchi, M., & Kitching, R. L. (2016). Identifying indicator species of elevation: Comparing the utility of woody plants, ants and moths for long- term monitoring. *Austral Ecology*, 41(2), 179-188.
- Nally, R. M., & Fleishman, E. (2004). A successful predictive model of species richness based on indicator species. *Conservation biology*, 18(3), 646-654.
- Nascimento, C. W. A. d., Biondi, C. M., Silva, F. B. V. d., & Lima, L. H. V. (2021). Using plants to remediate or manage metal-polluted soils: An overview on the current state of phytotechnologies. *Acta Scientiarum. Agronomy*, 43.
- Nasrullah, R. N., Bibi, H., Iqbal, M., & Durrani, M. I. (2006). Pollution load in industrial effluent and ground water of Gadoon Amazai Industrial Estate (GAIE) Swabi, NWFP. *Journal of agricultural and biological science*, 1(3), 18-24.
- Natasha, N., Shahid, M., Bibi, I., Iqbal, J., Khalid, S., Murtaza, B., . . . Hammad, H. M. (2022). Zinc in soil-plant-human system: A data-analysis review. *Science of the Total Environment*, 808, 152024.
- Nathan, N. S., Saravanane, R., & Sundararajan, T. (2017). Application of ANN and MLR models on groundwater quality using CWQI at Lawspet, Puducherry in India. *Journal of Geoscience and Environment Protection*, 5(03), 99.
- Nations, U. (2015). Department of Economic and Social Affairs Sustainable Development. The 17 Goals.
- Naveh, Z., & Whittaker, R. H. (1980). Structural and floristic diversity of shrublands and woodlands in northern Israel and other Mediterranean areas. *Vegetatio*, 41, 171-190.
- Nazir, A., Malik, R. N., Ajaib, M., Khan, N., & Siddiqui, M. F. (2011). Hyperaccumulators of heavy metals of industrial areas of Islamabad and Rawalpindi. *Pak J Bot*, 43(4), 1925-1933.
- Nazir, R., Khan, M., Masab, M., Rehman, H. U., Rauf, N. U., Shahab, S., . . . Rafeeq, M. (2015). Accumulation of heavy metals (Ni, Cu, Cd, Cr, Pb, Zn, Fe) in the soil, water and plants and analysis of physico-chemical parameters of soil and water collected from Tanda Dam Kohat. *Journal of pharmaceutical sciences and research*, 7(3), 89.
- Nguyen, K. T., Ahmed, M. B., Mojiri, A., Huang, Y., Zhou, J. L., & Li, D. (2021). Advances in As contamination and adsorption in soil for effective management. *Journal of Environmental Management*, 296, 113274.



- Nissim, W. G., & Labrecque, M. (2021). Reclamation of urban brownfields through phytoremediation: Implications for building sustainable and resilient towns. *Urban Forestry & Urban Greening*, 65, 127364.
- Nithamathi, C., & Indira, V. (2005). Impact of air pollution on *Cesalpinia sepiaria* Linn. *Tuticorin city. Indian Journal of Environment and Ecoplanning*, 10(2), 449-452.
- Njoku, E. A., & Tenenbaum, D. E. (2022). Quantitative assessment of the relationship between land use/land cover (LULC), topographic elevation and land surface temperature (LST) in Ilorin, Nigeria. *Remote Sensing Applications: Society and Environment*, 27, 100780.
- Nong, H., Liu, J., Chen, J., Zhao, Y., Wu, L., Tang, Y., . . . Xu, Z. (2023). Woody plants have the advantages in the phytoremediation process of manganese ore with the help of microorganisms. *Science of The Total Environment*, 863, 160995.
- Noor, M. J., Sultana, S., Fatima, S., Ahmad, M., Zafar, M., Sarfraz, M., . . . Ashraf, M. A. (2015). Retracted Article: Estimation of anticipated performance index and air pollution tolerance index and of vegetation around the marble industrial areas of Potwar region: bioindicators of plant pollution response. *Environmental Geochemistry and Health*, 37, 441-455.
- Noreen, I., Khan, S. M., Ahmad, Z., Rahman, I., Tabassum, A., & Abd\_Allah, E. (2019). Response of different plant species to pollution emitted from oil and gas plant with special reference to heavy metals accumulation. *Pak. J. Bot*, 51(4), 1231-1240.
- Notten, M., Oosthoek, A., Rozema, J., & Aerts, R. (2005). Heavy metal concentrations in a soil–plant–snail food chain along a terrestrial soil pollution gradient. *Environmental pollution*, 138(1), 178-190.
- Nouri, J., Khorasani, N., Lorestani, B., Karami, M., Hassani, A., & Yousefi, N. (2009). Accumulation of heavy metals in soil and uptake by plant species with phytoremediation potential. *Environmental Earth Sciences*, 59, 315-323.
- Nur-E-Alam, M., Mia, M. A. S., Ahmad, F., & Rahman, M. M. (2020). An overview of chromium removal techniques from tannery effluent. *Applied Water Science*, 10(9), 205.
- Odum, E. P. (1979). *Ecological Importance of the Riparian Zone*<sup>1</sup> 2. Paper presented at the Strategies for Protection and Management of Floodplain Wetlands and Other Riparian Ecosystems: Proceedings of the Symposium, December 11-13, 1978, Callaway Gardens, Georgia.
- Ogundele, O. D., & Anaun, T. E. (2022). Phytoremediation: A Green Approach for Pollution Cleanup.
- Oladoye, P. O., Olowe, O. M., & Asemoloye, M. D. J. C. (2022). Phytoremediation technology and food security impacts of heavy metal contaminated soils: A review of literature. 288, 132555.
- Olawoyin, R., Nieto, A., Grayson, R. L., Hardisty, F., & Oyewole, S. (2013). Application of artificial neural network (ANN)–self-organizing map (SOM) for the categorization of water, soil and sediment quality in petrochemical regions. *Expert Systems with Applications*, 40(9), 3634-3648.
- Organization, W. H. (1996). Permissible limits of heavy metals in soil and plants. *Geneva, Switzerland*.
- Organization, W. H. (2005). *The World Health Report 2005: Make every mother and child count*: World Health Organization.
- Organization, W. H. (2020). Chromium in Drinking-water: World Health Organization.
- Organization, W. H. (2021). Nickel in drinking water: background document for development of WHO Guidelines for drinking-water quality: World Health Organization.



- Organization, W. H. (2023). Trends in maternal mortality 2000 to 2020: estimates by WHO, UNICEF, UNFPA, World Bank Group and UNDESA/Population Division: executive summary.
- Oyewole, O. A., Zobeashia, S. S. L.-T., Oladoja, E. O., Raji, R. O., Odiniya, E. E., & Musa, A. M. (2019). Biosorption of heavy metal polluted soil using bacteria and fungi isolated from soil. *SN Applied Sciences*, 1, 1-8.
- Ozenen Kavlak, M., Cabuk, S. N., & Cetin, M. (2021). Development of forest fire risk map using geographical information systems and remote sensing capabilities: Ören case. *Environmental Science and Pollution Research*, 28(25), 33265-33291.
- Pakistan, W. (2007). From water and health related issues in Pakistan. *Fresh water and toxic programme*, 1-20.
- Pandian, S., Rakkammal, K., Rathinapriya, P., Rency, A. S., Satish, L., & Ramesh, M. (2020). Physiological and biochemical changes in sorghum under combined heavy metal stress: An adaptive defence against oxidative stress. *Biocatalysis and Agricultural Biotechnology*, 29, 101830.
- Papazoglou, E. G., Serelis, K. G., & Bouranis, D. L. J. E. J. o. S. B. (2007). Impact of high cadmium and nickel soil concentration on selected physiological parameters of *Arundo donax* L. 43(4), 207-215.
- Parmar, P., Dave, B., Sudhir, A., Panchal, K., & Subramanian, R. (2013). Physiological, biochemical and molecular response of plants against heavy metals stress. *International Journal of Current Research*, 5(1), 80-89.
- Parween, S., Siddique, N. A., Diganta, M. T. M., Olbert, A. I., & Uddin, M. G. (2022). Assessment of urban river water quality using modified NSF water quality index model at Siliguri city, West Bengal, India. *Environmental and Sustainability Indicators*, 16, 100202.
- Paudel, B., Zhang, Y.-l., Li, S.-c., Liu, L.-s., Wu, X., & Khanal, N. R. (2016). Review of studies on land use and land cover change in Nepal. *Journal of Mountain Science*, 13, 643-660.
- Pérez- Escobar, O. A., Chomicki, G., Condamine, F. L., Karremans, A. P., Bogarín, D., Matzke, N. J., . . . Antonelli, A. (2017). Recent origin and rapid speciation of Neotropical orchids in the world's richest plant biodiversity hotspot. *New Phytologist*, 215(2), 891-905.
- Piotto, F. A., Carvalho, M. E. A., Souza, L. A., Rabêlo, F. H. S., Franco, M. R., Batagin-Piotto, K. D., . . . Research, P. (2018). Estimating tomato tolerance to heavy metal toxicity: cadmium as study case. 25, 27535-27544.
- Polikar, R. (2012). Ensemble learning. *Ensemble machine learning: Methods and applications*, 1-34.
- Poonam, T., Tanushree, B., & Sukalyan, C. (2013). Water quality indices-important tools for water quality assessment: a review. *International Journal of Advances in chemistry*, 1(1), 15-28.
- Pouresmaeli, M., Ataei, M., Forouzandeh, P., Azizollahi, P., & Mahmoudifard, M. (2022). Recent progress on sustainable phytoremediation of heavy metals from soil. *Journal of Environmental Chemical Engineering*, 108482.
- Priya, A., Muruganandam, M., Ali, S. S., & Kornaros, M. J. T. (2023). Clean-Up of Heavy Metals from Contaminated Soil by Phytoremediation: A Multidisciplinary and Eco-Friendly Approach. 11(5), 422.
- Puckett, L. J. (1995). Identifying the major sources of nutrient water pollution. *Environmental Science & Technology*, 29(9), 408A-414A.



- Qadir, A., & Malik, R. N. (2009). Assessment of an index of biological integrity (IBI) to quantify the quality of two tributaries of river Chenab, Sialkot, Pakistan. *Hydrobiologia*, 621, 127-153.
- Qadir, A., Malik, R. N., & Husain, S. Z. (2008a). Spatio-temporal variations in water quality of Nullah Aik-tributary of the river Chenab, Pakistan. *Environmental monitoring and assessment*, 140, 43-59.
- Qadir, A., Malik, R. N., & Husain, S. Z. (2008b). Spatio-temporal variations in water quality of Nullah Aik-tributary of the river Chenab, Pakistan. *Environmental Monitoring and Assessment*, 140(1), 43-59. doi: 10.1007/s10661-007-9846-4
- Qadir, S., Hameed, A., Nisa, N., Azooz, M., Wani, M. R., Hasannuzaman, M., . . . Ahmad, P. J. I. o. C. i. t. E. o. C. C. V. (2014). Brassicas: responses and tolerance to heavy metal stress. 1-36.
- Qing, X., Yutong, Z., & Shenggao, L. (2015). Assessment of heavy metal pollution and human health risk in urban soils of steel industrial city (Anshan), Liaoning, Northeast China. *Ecotoxicology and environmental safety*, 120, 377-385.
- Qu, S., Shi, Z., Liang, X., Wang, G., & Han, J. (2021). Multiple factors control groundwater chemistry and quality of multi-layer groundwater system in Northwest China coalfield—Using self-organizing maps (SOM). *Journal of Geochemical Exploration*, 227, 106795.
- Quinteros, E., Ribó, A., Mejía, R., López, A., Belteton, W., Comandari, A., . . . López, D. L. (2017). Heavy metals and pesticide exposure from agricultural activities and former agrochemical factory in a Salvadoran rural community. *Environmental Science and Pollution Research*, 24, 1662-1676.
- Rabelo, L. M., Guimarães, A. T. B., de Souza, J. M., da Silva, W. A. M., de Oliveira Mendes, B., de Oliveira Ferreira, R., . . . Malafaia, G. (2018). Correction to: histological liver changes in Swiss mice caused by tannery effluent. *Environmental science and pollution research international*, 25(16), 16267-16268.
- Rahman, A. u., Khan, S. M., Ahmad, Z., Alamri, S., Hashem, M., Ilyas, M., . . . Shahab Ali, G. K. (2021). -Impact of multiple environmental factors on species abundance in various forest layers using an integrative modeling approach. *Global Ecology and Conservation*, 29, e01712. doi: <https://doi.org/10.1016/j.gecco.2021.e01712>
- Rahman, Z., & Singh, V. P. (2019). The relative impact of toxic heavy metals (THMs)(arsenic (As), cadmium (Cd), chromium (Cr)(VI), mercury (Hg), and lead (Pb)) on the total environment: an overview. *Environmental monitoring and assessment*, 191, 1-21.
- Ramakrishnaiah, C., Sadashivaiah, C., & Ranganna, G. (2009). Assessment of water quality index for the groundwater in Tumkur Taluk, Karnataka State, India. *E-Journal of chemistry*, 6(2), 523-530.
- Rani, P., Arya, R. C., & Dwivedi, S. (2020). Chromium pollution: Impact on plants and its mitigation. *Innovations in Food Technology: Current Perspectives and Future Goals*, 323-340.
- Ranieri, E., & Gikas, P. (2014). Effects of plants for reduction and removal of hexavalent chromium from a contaminated soil. *Water, air, & soil pollution*, 225, 1-9.
- Ranjith, S. (2019). Utilization of Water Quality Modeling and Dissolved Oxygen Control in River Tungabhadra, Karnataka (India). *Open Access Library Journal*, 6(05), 1.
- Raoof Mahmood, A., Alheety, M. A., Asker, M. M., Zyaad Tareq, A., & Karadağ, A. (2019). *Saccharine based carbonyl multi-walled carbon nanotubes: novel modification, characterization and its ability for removing Cd (II) and Cu (II) from soil and environmental water samples*. Paper presented at the Journal of Physics: Conference Series.



- Rashid, A., Ayub, M., Khan, S., Ullah, Z., Ali, L., Gao, X., . . . Rasool, A. (2022). Hydrogeochemical assessment of carcinogenic and non-carcinogenic health risks of potentially toxic elements in aquifers of the Hindukush ranges, Pakistan: Insights from groundwater pollution indexing, GIS-based, and multivariate statistical approaches. *Environmental Science and Pollution Research*, 29(50), 75744-75768.
- Raskin, I., Kumar, P. N., Dushenkov, S., & Salt, D. E. (1994). Bioconcentration of heavy metals by plants. *Current Opinion in biotechnology*, 5(3), 285-290.
- Ravankhah, N., Mirzaei, R., & Masoum, S. (2015). Evaluation of geoaccumulation index, contamination factor, and principal component analysis for estimating soil contamination. *Iranian journal of health and environment*, 8(3).
- Reed, B. M., Sarasan, V., Kane, M., Bunn, E., & Pence, V. C. (2011). Biodiversity conservation and conservation biotechnology tools. *In Vitro Cellular & Developmental Biology-Plant*, 47, 1-4.
- Revankar, N., & Kadadevaru, G. G. (2021). Physico-chemical Parameters and Zooplankton Community at Sangave Pond in Uttara Kannada District, Karnataka, India. *Environmental Sciences*, 9(12), 998-1003.
- Rizwan, M., Ali, S., Zaheer Akbar, M., Shakoor, M. B., Mahmood, A., Ishaque, W., . . . Research, P. (2017). Foliar application of aspartic acid lowers cadmium uptake and Cd-induced oxidative stress in rice under Cd stress. 24, 21938-21947.
- Robinson, B., Anderson, C., & Dickinson, N. (2015). Phytoextraction: where's the action? *Journal of Geochemical Exploration*, 151, 34-40.
- Rockström, J., Falkenmark, M., Allan, T., Folke, C., Gordon, L., Jägerskog, A., . . . Molden, D. (2014). The unfolding water drama in the Anthropocene: towards a resilience-based perspective on water for global sustainability. *Ecohydrology*, 7(5), 1249-1261.
- Rodier, J., Bazin, C., Broutin, J., Chambon, P., Champsaur, H., & Rodi, L. (2009). Water analysis, 9th edit. *Dunod, Paris, France*, 1579.
- Rodrigues, S. M., & Römkens, P. F. (2018). Human health risks and soil pollution *Soil Pollution* (pp. 217-250): Elsevier.
- Rondinelli, D. A., Johnson Jr, J. H., & Kasarda, J. D. (1998). The changing forces of urban economic development: Globalization and city competitiveness in the 21st century. *Cityscape*, 71-105.
- Rosseel, Y., Oberski, D., Byrnes, J., Vanbrabant, L., Savalei, V., Merkle, E., . . . Barendse, M. (2017). Package 'lavaan'. Retrieved June, 17(1), 2017.
- Rowe, R. J. (2009). Environmental and geometric drivers of small mammal diversity along elevational gradients in Utah. *Ecography*, 32(3), 411-422.
- Saarela, J. M., Burke, S. V., Wysocki, W. P., Barrett, M. D., Clark, L. G., Craine, J. M., . . . Duvall, M. R. (2018). A 250 plastome phylogeny of the grass family (Poaceae): topological support under different data partitions. *PeerJ*, 6, e4299.
- Sabins Jr, F. F., & Ellis, J. M. (2020). *Remote sensing: Principles, interpretation, and applications*: Waveland Press.
- Sadiq, Y., Zaid, A., & Khan, M. M. A. (2020). Adaptive physiological responses of plants under abiotic stresses: role of phytohormones. *Plant Ecophysiology and Adaptation under Climate Change: Mechanisms and Perspectives I: General Consequences and Plant Responses*, 797-824.
- Saha, J. K., Selladurai, R., Coumar, M. V., Dotaniya, M., Kundu, S., & Patra, A. K. (2017). *Soil pollution-an emerging threat to agriculture*: Springer.
- Said, I., Salman, S. A., & Elnazer, A. A. (2019). Multivariate statistics and contamination factor to identify trace elements pollution in soil around Gerga City, Egypt. *Bulletin of the National Research Centre*, 43, 1-6.



- Sakaa, B., Elbeltagi, A., Boudibi, S., Chaffai, H., Islam, A. R. M. T., Kulimushi, L. C., . . . Wong, Y. J. (2022). Water quality index modeling using random forest and improved SMO algorithm for support vector machine in Saf-Saf river basin. *Environmental Science and Pollution Research*, 29(32), 48491-48508.
- Sakizadeh, M. (2015). Assessment the performance of classification methods in water quality studies, A case study in Karaj River. *Environmental monitoring and assessment*, 187, 1-12.
- Saleem, M. H., Ali, S., Rehman, M., Rizwan, M., Kamran, M., Mohamed, I. A., . . . Liu, L. (2020). Individual and combined application of EDTA and citric acid assisted phytoextraction of copper using jute (*Corchorus capsularis* L.) seedlings. *Environmental Technology & Innovation*, 19, 100895.
- Samudro, G., & Mangkoedihardjo, S. (2020). Mixed plant operations for phytoremediation in polluted environments—a critical review. *Journal of Phytology*, 12, 99-103.
- Sankhla, M. S., Kumar, R., & Prasad, L. (2019). Distribution and contamination assessment of potentially harmful element chromium in water. *Available at SSRN 3492307*.
- Saradhi, P. P. (1991). Proline accumulation under heavy metal stress. *Journal of Plant Physiology*, 138(5), 554-558.
- Sarfraz, W., Farid, M., Ejaz, U., Naeem, N., Yousaf, A., Kabir, Z., . . . Khalid, N. (2022a). EDTA could enhance the bioaccumulation of Pb and Cd in *Ricinus communis* for potential decontamination of roadside soils. *Fresenius Environ Bull*, 31(01), 81-91.
- Sarfraz, W., Farid, M., Ejaz, U., Naeem, N., Yousaf, A., Kabir, Z., . . . Khalid, N. J. F. E. B. (2022b). EDTA could enhance the bioaccumulation of Pb and Cd in *Ricinus communis* for potential decontamination of roadside soils. 31(01), 81-91.
- Sarma, H. (2011). Metal hyperaccumulation in plants: a review focusing on phytoremediation technology. *Journal of Environmental Science and Technology*, 4(2), 118-138.
- Saxena, G., Purchase, D., Mulla, S. I., Saratale, G. D., & Bharagava, R. N. (2020). Phytoremediation of heavy metal-contaminated sites: eco-environmental concerns, field studies, sustainability issues, and future prospects. *Reviews of Environmental Contamination and Toxicology Volume 249*, 71-131.
- Schickler, H., & Caspi, H. (1999). Response of antioxidative enzymes to nickel and cadmium stress in hyperaccumulator plants of the genus *Alyssum*. *Physiologia plantarum*, 105(1), 39-44.
- Schiff, R., & Benoit, G. (2007). Effects of impervious cover at multiple spatial scales on coastal watershed streams 1. *JAWRA Journal of the American Water Resources Association*, 43(3), 712-730.
- Shahid, M. A., Balal, R. M., Pervez, M. A., Abbas, T., Aqeel, M. A., JAVAID, M., & Garcia-Sanchez, F. (2014). Exogenous proline and proline-enriched *Lolium perenne* leaf extract protects against phytotoxic effects of nickel and salinity in *Pisum sativum* by altering polyamine metabolism in leaves. *Turkish Journal of Botany*, 38(5), 914-926.
- Sharma, M. R., & Raju, N. (2013). Correlation of heavy metal contamination with soil properties of industrial areas of Mysore, Karnataka, India by cluster analysis. *International Research Journal of Environment Sciences*, 2(10), 22-27.
- Shen, X., Dai, M., Yang, J., Sun, L., Tan, X., Peng, C., . . . Naz, I. (2022). A critical review on the phytoremediation of heavy metals from environment: Performance and challenges. *Chemosphere*, 291, 132979.
- Shrestha, R., Ban, S., Devkota, S., Sharma, S., Joshi, R., Tiwari, A. P., . . . Joshi, M. K. (2021). Technological trends in heavy metals removal from industrial wastewater: A review. *Journal of Environmental Chemical Engineering*, 9(4), 105688.
- Shrestha, S., Bastola, S., Babel, M., Dulal, K., Magome, J., Hapuarachchi, H., . . . Takeuchi, K. (2007). The assessment of spatial and temporal transferability of a physically based



- distributed hydrological model parameters in different physiographic regions of Nepal. *Journal of Hydrology*, 347(1-2), 153-172.
- Shrivastava, A., Ghosh, D., Dash, A., & Bose, S. (2015). Arsenic contamination in soil and sediment in India: sources, effects, and remediation. *Current Pollution Reports*, 1, 35-46.
- Sial, R., Chaudhary, M., Abbas, S., Latif, M., & Khan, A. (2006). Quality of effluents from Hattar industrial estate. *Journal of Zhejiang University SCIENCE B*, 7, 974-980.
- Siddique, A., Kandpal, G., Kumar, P. J. J. o. P., & Microbiology, A. (2018). Proline accumulation and its defensive role under diverse stress condition in plants: An overview. *12*(3), 1655-1659.
- Simonova, O., & Cheglacova, O. (2017). Influence of fertilizers on content and dynamics of mobile forms of copper and zinc in sod-podzolic soil. *Agricultural Science Euro-North-East*(6), 30-34.
- Singh, H., & Pant, G. (2023). Phytoremediation: Low input-based ecological approach for sustainable environment. *Applied Water Science*, 13(3), 85. doi: 10.1007/s13201-023-01898-2
- Singh, K. P., Malik, A., Sinha, S., Singh, V. K., & Murthy, R. C. (2005). Estimation of source of heavy metal contamination in sediments of Gomti River (India) using principal component analysis. *Water, air, and soil pollution*, 166, 321-341.
- Singh, K. R., Goswami, A. P., Kalamdhad, A. S., & Kumar, B. (2020). Assessment of surface water quality of Pagladia, Beki and Kolong rivers (Assam, India) using multivariate statistical techniques. *International Journal of River Basin Management*, 18(4), 511-520.
- Singh, M., Singh, V. P., Dubey, G., & Prasad, S. M. (2015). Exogenous proline application ameliorates toxic effects of arsenate in *Solanum melongena* L. seedlings. *Ecotoxicology and Environmental Safety*, 117, 164-173.
- Singha, S., Pasupuleti, S., Singha, S. S., Singh, R., & Kumar, S. (2021). Prediction of groundwater quality using efficient machine learning technique. *Chemosphere*, 276, 130265.
- Sinha, P., Verma, N. K., & Ayele, E. (2016). Urban built-up area extraction and change detection of Adama municipal area using time-series Landsat images. *Int. J. Adv. Remote Sens. GIS*, 5(8), 1886-1895.
- Siraj, M., Murtaza, B. N., Sardar, A., Muntaha, S. T., Ali, P. A., & Chivers, D. (2022). Toxicological impact of water pollutants on DNA and tissues of inhabitant fish *Labeo dyocheilus* of River Kabul, Khyber Pakhtunkhwa, Pakistan. *Natural and Applied Sciences International Journal (NASIJ)*, 3(2), 85-99.
- Siripornadulsil, S., Traina, S., Verma, D. P. S., & Sayre, R. T. J. T. P. C. (2002). Molecular mechanisms of proline-mediated tolerance to toxic heavy metals in transgenic microalgae. *14*(11), 2837-2847.
- Skidmore, A. K., Bijker, W., Schmidt, K., & Kumar, L. (1997). Use of remote sensing and GIS for sustainable land management. *ITC journal*, 3(4), 302-315.
- Soares, C., de Sousa, A., Pinto, A., Azenha, M., Teixeira, J., Azevedo, R. A., . . . Botany, E. (2016). Effect of 24-epibrassinolide on ROS content, antioxidant system, lipid peroxidation and Ni uptake in *Solanum nigrum* L. under Ni stress. *122*, 115-125.
- Song, B., Zeng, G., Gong, J., Liang, J., Xu, P., Liu, Z., . . . Liu, Y. (2017). Evaluation methods for assessing effectiveness of in situ remediation of soil and sediment contaminated with organic pollutants and heavy metals. *Environment international*, 105, 43-55.



- Sonone, S. S., Jadhav, S., Sankhla, M. S., & Kumar, R. (2020). Water contamination by heavy metals and their toxic effect on aquaculture and human health through food Chain. *Lett. Appl. NanoBioScience*, 10(2), 2148-2166.
- Stohlgren, T. J., Bull, K. A., & Otsuki, Y. (1998). Comparison of rangeland vegetation sampling techniques in the Central Grasslands.
- Sumiahadi, A., & Acar, R. (2018). *A review of phytoremediation technology: heavy metals uptake by plants*. Paper presented at the IOP conference series: earth and environmental science.
- Sun, A. Y., & Scanlon, B. R. (2019). How can Big Data and machine learning benefit environment and water management: a survey of methods, applications, and future directions. *Environmental Research Letters*, 14(7), 073001.
- Suresh, G., Ramasamy, V., Meenakshisundaram, V., Venkatachalapathy, R., & Ponnusamy, V. (2011). Influence of mineralogical and heavy metal composition on natural radionuclide concentrations in the river sediments. *Applied Radiation and Isotopes*, 69(10), 1466-1474. doi: <https://doi.org/10.1016/j.apradiso.2011.05.020>
- Syfert, M. M., Smith, M. J., & Coomes, D. A. (2013). The effects of sampling bias and model complexity on the predictive performance of MaxEnt species distribution models. *PloS one*, 8(2), e55158.
- Szálkai, K. (2023). Water-borne diseases *The Palgrave Encyclopedia of Global Security Studies* (pp. 1540-1546): Springer.
- Tadesse, G. L., Guya, T. K., & Walabu, M. (2017). Impacts of tannery effluent on environments and human health: a review article. *Advances in Life Science and Technology*, 54(10).
- Tahkokorpi, M., Taulavuori, K., Laine, K., & Taulavuori, E. (2007). After-effects of drought-related winter stress in previous and current year stems of *Vaccinium myrtillus* L. *Environmental and experimental botany*, 61(1), 85-93.
- Tang, B., Xu, H., Song, F., Ge, H., & Yue, S. (2022). Effects of heavy metals on microorganisms and enzymes in soils of lead–zinc tailing ponds. *Environmental Research*, 207, 112174.
- Tang, W.-L., Liu, Y.-R., Guan, W.-Y., Zhong, H., Qu, X.-M., & Zhang, T. (2020). Understanding mercury methylation in the changing environment: Recent advances in assessing microbial methylators and mercury bioavailability. *Science of the Total Environment*, 714, 136827.
- Tangahu, B. V., Sheikh Abdullah, S. R., Basri, H., Idris, M., Anuar, N., & Mukhlisin, M. (2011). A review on heavy metals (As, Pb, and Hg) uptake by plants through phytoremediation. *International journal of chemical engineering*, 2011.
- Tariq, A., & Mumtaz, F. (2023). Modeling spatio-temporal assessment of land use land cover of Lahore and its impact on land surface temperature using multi-spectral remote sensing data. *Environmental Science and Pollution Research*, 30(9), 23908-23924.
- Tariq, A., & Mushtaq, A. (2023). Untreated wastewater reasons and causes: a review of most affected areas and cities. *Int. J. Chem. Biochem. Sci*, 23, 121-143.
- Tariq, S. R., Shaheen, N., Khalique, A., & Shah, M. H. (2010). Distribution, correlation, and source apportionment of selected metals in tannery effluents, related soils, and groundwater—a case study from Multan, Pakistan. *Environmental monitoring and assessment*, 166, 303-312.
- Teng, D., Mao, K., Ali, W., Xu, G., Huang, G., Niazi, N. K., . . . Zhang, H. (2020). Describing the toxicity and sources and the remediation technologies for mercury-contaminated soil. *RSC advances*, 10(39), 23221-23232.
- Ter Braak, C. J., & Prentice, I. C. (1988). A theory of gradient analysis *Advances in ecological research* (Vol. 18, pp. 271-317): Elsevier.



- Thakur, M., Praveen, S., Divte, P. R., Mitra, R., Kumar, M., Gupta, C. K., . . . Anand, A. (2022). Metal tolerance in plants: Molecular and physicochemical interface determines the “not so heavy effect” of heavy metals. *Chemosphere*, 287, 131957.
- Tian, K., Huang, B., Xing, Z., & Hu, W. (2017). Geochemical baseline establishment and ecological risk evaluation of heavy metals in greenhouse soils from Dongtai, China. *Ecological Indicators*, 72, 510-520. doi: <https://doi.org/10.1016/j.ecolind.2016.08.037>
- Tian, K., Wu, Q., Liu, P., Hu, W., Huang, B., Shi, B., . . . Ryu, J. (2020). Ecological risk assessment of heavy metals in sediments and water from the coastal areas of the Bohai Sea and the Yellow Sea. *Environment international*, 136, 105512.
- Tomaszewski, B. (2020). *Geographic information systems (GIS) for disaster management*: Routledge.
- Tu, S., & Ma, L. (2003). Interactive effects of pH, arsenic and phosphorus on uptake of As and P and growth of the arsenic hyperaccumulator *Pteris vittata* L. under hydroponic conditions. *Environmental and Experimental Botany*, 50(3), 243-251.
- Uddin, M. G., Nash, S., Rahman, A., & Olbert, A. I. (2022). A comprehensive method for improvement of water quality index (WQI) models for coastal water quality assessment. *Water Research*, 219, 118532.
- Ukaogo, P. O., Ewuzie, U., & Onwuka, C. V. (2020). Environmental pollution: causes, effects, and the remedies *Microorganisms for sustainable environment and health* (pp. 419-429): Elsevier.
- Ullah, A., Al-Busaidi, W. M., Al-Sadi, A. M., & Farooq, M. (2022). Bread wheat genotypes accumulating free proline and phenolics can better tolerate drought stress through sustained rate of photosynthesis. *Journal of Soil Science and Plant Nutrition*, 22(1), 165-176.
- Unesco, I. (2020). Basic texts of the 2003 convention for the safeguarding of the intangible cultural heritage.
- Usman, K., Al-Ghouti, M. A., & Abu-Dieyeh, M. H. (2018). Phytoremediation: halophytes as promising heavy metal hyperaccumulators. *Heavy Met*, 27, 7378.
- Usman, K., Al Jabri, H., Abu-Dieyeh, M. H., & Alsafran, M. H. (2020). Comparative assessment of toxic metals bioaccumulation and the mechanisms of chromium (Cr) tolerance and uptake in *Calotropis procera*. *Frontiers in plant science*, 11, 883.
- Utz, R. M., Hopkins, K. G., Beesley, L., Booth, D. B., Hawley, R. J., Baker, M. E., . . . L. Jones, K. (2016). Ecological resistance in urban streams: the role of natural and legacy attributes. *Freshwater Science*, 35(1), 380-397.
- Vaiopoulou, E., & Gikas, P. (2020). Regulations for chromium emissions to the aquatic environment in Europe and elsewhere. *Chemosphere*, 254, 126876.
- Van Groenewoud, H. (1992). The robustness of correspondence, detrended correspondence, and TWINSpan analysis. *Journal of Vegetation Science*, 3(2), 239-246.
- Van Straalen, N. (1996). 4 Uptake of pollutants by soil organisms. *Soil Pollution and Soil Protection*, 55.
- Vara Prasad, M. N., & de Oliveira Freitas, H. M. (2003). Metal hyperaccumulation in plants: biodiversity prospecting for phytoremediation technology. *Electronic journal of biotechnology*, 6(3), 285-321.
- Vesanto, J., & Alhoniemi, E. (2000). Clustering of the self-organizing map. *IEEE Transactions on neural networks*, 11(3), 586-600.
- Vesanto, J., Himberg, J., Alhoniemi, E., Parhankangas, J., Team, S., & Oy, L. (2000). SOM toolbox for Matlab 5: Citeseer.
- Vodyanitskii, Y. N. J. E. S. S. (2013). Contamination of soils with heavy metals and metalloids and its ecological hazard (analytic review). 46, 793-801.



- Vorobeichik, E. (2022). Natural recovery of terrestrial ecosystems after the cessation of industrial pollution: 1. A state-of-the-art review. *Russian Journal of Ecology*, 53(1), 1-39.
- Wang, J., Bretz, M., Dewan, M. A. A., & Delavar, M. A. (2022). Machine learning in modelling land-use and land cover-change (LULCC): Current status, challenges and prospects. *Science of the Total Environment*, 822, 153559.
- Wang, Q., Jiang, D., Gao, Y., Zhang, Z., & Chang, Q. (2022). Examining the Driving Factors of SOM Using a Multi-Scale GWR Model Augmented by Geo-Detector and GWPCA Analysis. *Agronomy*, 12(7), 1697.
- Wang, X. (2020). Urbanization Trends and Challenges in Developing Countries. *International Journal of Business Management and Visuals*, ISSN: 3006-2705, 3(1), 8-14.
- Wang, Y., Tan, P., Chang, L., Yue, Z., Zeng, C., Li, M., . . . Yan, M. (2022). Exogenous proline mitigates toxic effects of cadmium via the decrease of cadmium accumulation and reestablishment of redox homeostasis in *Brassica juncea*. *BMC Plant Biology*, 22(1), 1-19.
- Wani, W., Masoodi, K. Z., Zaid, A., Wani, S. H., Shah, F., Meena, V. S., . . . Mosa, K. A. J. R. L. S. F. e. N. (2018). Engineering plants for heavy metal stress tolerance. 29, 709-723.
- Wantzen, K. M., Alves, C. B. M., Badiane, S. D., Bala, R., Blettler, M., Callisto, M., . . . Leite, M. F. (2019). Urban stream and wetland restoration in the Global South—A DPSIR analysis. *Sustainability*, 11(18), 4975.
- Water, U. (2017). Wastewater—The Untapped Resource; The United Nations World Water Development Report 2017: UNESCO.
- Wei, B., Jiang, F., Li, X., & Mu, S. (2010). Heavy metal induced ecological risk in the city of Urumqi, NW China. *Environmental monitoring and assessment*, 160, 33-45.
- Wellmann, T., Lausch, A., Andersson, E., Knapp, S., Cortinovis, C., Jache, J., . . . Kraemer, R. (2020). Remote sensing in urban planning: Contributions towards ecologically sound policies? *Landscape and urban planning*, 204, 103921.
- Whitehead, P., Bussi, G., Hossain, M. A., Dolk, M., Das, P., Comber, S., . . . Hossain, M. S. (2018). Restoring water quality in the polluted Turag-Tongi-Balu river system, Dhaka: Modelling nutrient and total coliform intervention strategies. *Science of the total environment*, 631, 223-232.
- Whittaker, R. H. (1978). Direct gradient analysis *Ordination of plant communities* (pp. 7-50): Springer.
- Wimalawansa, S. J. (2016). The role of ions, heavy metals, fluoride, and agrochemicals: critical evaluation of potential aetiological factors of chronic kidney disease of multifactorial origin (CKDmfo/CKDu) and recommendations for its eradication. *Environmental geochemistry and health*, 38(3), 639-678.
- Wong, W. Y., Al-Ani, A. K. I., Hasikin, K., Khairuddin, A. S. M., Razak, S. A., Hizaddin, H. F., . . . Azizan, M. M. (2022). Water quality index using modified random forest technique: assessing novel input features. *CMES-Computer Modeling in Engineering & Sciences*.
- Wu, H., & Li, Z.-L. (2009). Scale issues in remote sensing: A review on analysis, processing and modeling. *Sensors*, 9(3), 1768-1793.
- Wu, Y., Li, S., & Yu, S. (2016). Monitoring urban expansion and its effects on land use and land cover changes in Guangzhou city, China. *Environmental monitoring and assessment*, 188, 1-15.



- Xiang, Q., Yu, H., Chu, H., Hu, M., Xu, T., Xu, X., & He, Z. (2022). The potential ecological risk assessment of soil heavy metals using self-organizing map. *Science of the Total Environment*, 843, 156978.
- Xie, H., Zhang, Y., Zeng, X., & He, Y. (2020). Sustainable land use and management research: A scientometric review. *Landscape Ecology*, 35, 2381-2411.
- Xu, D.-M., Yan, B., Chen, T., Lei, C., Lin, H.-Z., & Xiao, X.-M. (2017). Contaminant characteristics and environmental risk assessment of heavy metals in the paddy soils from lead (Pb)-zinc (Zn) mining areas in Guangdong Province, South China. *Environmental Science and Pollution Research*, 24, 24387-24399.
- Xu, J., Yin, H., & Li, X. (2009). Protective effects of proline against cadmium toxicity in micropropagated hyperaccumulator, *Solanum nigrum* L. *Plant cell reports*, 28, 325-333.
- Yaashikaa, P., Kumar, P. S., Jeevanantham, S., & Saravanan, R. (2022). A review on bioremediation approach for heavy metal detoxification and accumulation in plants. *Environmental Pollution*, 301, 119035.
- Yang, C., Jia, H., Dong, L., Zhao, H., & Zhao, M. (2024). Selection of Landsat-8 Operational Land Imager (OLI) Optimal Band Combinations for Mapping Alteration Zones. *Remote Sensing*, 16(2), 392.
- Yang, D., Yang, Y., & Xia, J. (2021). Hydrological cycle and water resources in a changing world: A review. *Geography and Sustainability*, 2(2), 115-122.
- Yang, S., Sun, L., Sun, Y., Song, K., Qin, Q., Zhu, Z., & Xue, Y. (2023). Towards an integrated health risk assessment framework of soil heavy metals pollution: Theoretical basis, conceptual model, and perspectives. *Environmental Pollution*, 316, 120596.
- Yang, X., Feng, Y., He, Z., & Stoffella, P. J. (2005). Molecular mechanisms of heavy metal hyperaccumulation and phytoremediation. *Journal of trace elements in medicine and biology*, 18(4), 339-353.
- Yang, Z., Yang, F., Liu, J.-L., Wu, H.-T., Yang, H., Shi, Y., . . . Chen, K.-M. (2022). Heavy metal transporters: Functional mechanisms, regulation, and application in phytoremediation. *Science of The Total Environment*, 809, 151099.
- Yarahmadi, S. S., & Ansari, M. R. (2018). Ecological risk assessment of heavy metals (Zn, Cr, Pb, As and Cu) in sediments of Dohezar River, North of Iran, Tonekabon city. *Acta Ecologica Sinica*, 38(2), 126-134.
- Yeh, G., Hoang, H.-G., Lin, C., Bui, X.-T., Tran, H.-T., Shern, C.-C., & Vu, C.-T. (2020). Assessment of heavy metal contamination and adverse biological effects of an industrially affected river. *Environmental Science and Pollution Research*, 27, 34770-34780.
- Yengoh, G. T., Dent, D., Olsson, L., Tengberg, A. E., & Tucker III, C. J. (2015). *Use of the Normalized Difference Vegetation Index (NDVI) to assess Land degradation at multiple scales: current status, future trends, and practical considerations*: Springer.
- Yevugah, L. L., Darko, G., & Bak, J. (2021). Does mercury emission from small-scale gold mining cause widespread soil pollution in Ghana? *Environmental Pollution*, 284, 116945.
- Yilma, M., Kiflie, Z., Windsperger, A., & Gessese, N. (2018). Application of artificial neural network in water quality index prediction: a case study in Little Akaki River, Addis Ababa, Ethiopia. *Modeling Earth Systems and Environment*, 4, 175-187.
- Yin, J., Yin, Z., Zhong, H., Xu, S., Hu, X., Wang, J., & Wu, J. (2011). Monitoring urban expansion and land use/land cover changes of Shanghai metropolitan area during the transitional economy (1979–2009) in China. *Environmental monitoring and assessment*, 177, 609-621.



- Zakaullah, & Ejaz, N. (2020). Investigation of the Soan River Water Quality Using Multivariate Statistical Approach. *International Journal of Photoenergy*, 2020, 1-15.
- Zamora-Ledezma, C., Negrete-Bolagay, D., Figueroa, F., Zamora-Ledezma, E., Ni, M., Alexis, F., & Guerrero, V. H. (2021). Heavy metal water pollution: A fresh look about hazards, novel and conventional remediation methods. *Environmental Technology & Innovation*, 22, 101504.
- Zehnder, A. J., Yang, H., & Schertenleib, R. (2003). Water issues: the need for action at different levels. *Aquatic sciences*, 65, 1-20.
- Zeng, F., Ali, S., Zhang, H., Ouyang, Y., Qiu, B., Wu, F., & Zhang, G. (2011). The influence of pH and organic matter content in paddy soil on heavy metal availability and their uptake by rice plants. *Environmental pollution*, 159(1), 84-91.
- Zeng, P., Guo, Z., Xiao, X., Peng, C., Liu, L., Yan, D., & He, Y. (2020). Physiological stress responses, mineral element uptake and phytoremediation potential of *Morus alba* L. in cadmium-contaminated soil. *Ecotoxicology and environmental safety*, 189, 109973.
- Zengin, F. K., Munzuroglu, O. J. A. S. S. B.-S., & Science, P. (2006). Toxic effects of cadmium (Cd<sup>++</sup>) on metabolism of sunflower (*Helianthus annuus* L.) seedlings. 56(3), 224-229.
- Zhang, C., & Li, X. (2022). Land Use and Land Cover Mapping in the Era of Big Data. *Land*, 11(10), 1692.
- Zhang, H., Yin, S., Chen, Y., Shao, S., Wu, J., Fan, M., . . . 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.
- Zhang, L., Liu, Z., Ren, T., Liu, D., Ma, Z., Tong, L., . . . Li, S. (2020). Identification of seed maize fields with high spatial resolution and multiple spectral remote sensing using random forest classifier. *Remote Sensing*, 12(3), 362.
- Zhang, Z., Tao, F., Du, J., Shi, P., Yu, D., Meng, Y., & Sun, Y. (2010). Surface water quality and its control in a river with intensive human impacts—a case study of the Xiangjiang River, China. *Journal of environmental management*, 91(12), 2483-2490.
- Zheng, Y., Tang, L., & Wang, H. (2021). An improved approach for monitoring urban built-up areas by combining NPP-VIIRS nighttime light, NDVI, NDWI, and NDBI. *Journal of Cleaner Production*, 328, 129488.
- Zhong, C., Guo, H., Swan, I., Gao, P., Yao, Q., & Li, H. (2023). Evaluating trends, profits, and risks of global cities in recent urban expansion for advancing sustainable development. *Habitat International*, 138, 102869.
- Zhong, X., Chen, Z., Li, Y., Ding, K., Liu, W., Liu, Y., . . . Yang, W. (2020). Factors influencing heavy metal availability and risk assessment of soils at typical metal mines in Eastern China. *Journal of hazardous materials*, 400, 123289.
- Zhou, X.-Y., & Wang, X.-R. (2019). Impact of industrial activities on heavy metal contamination in soils in three major urban agglomerations of China. *Journal of Cleaner Production*, 230, 1-10.
- Zhou, Z., Chen, K., Li, X., Zhang, S., Wu, Y., Zhou, Y., . . . Fan, W. (2020). Sign-to-speech translation using machine-learning-assisted stretchable sensor arrays. *Nature Electronics*, 3(9), 571-578.
- Zouari, M., Elloumi, N., Ahmed, C. B., Delmail, D., Rouina, B. B., Abdallah, F. B., & Labrousse, P. (2016). Exogenous proline enhances growth, mineral uptake, antioxidant defense, and reduces cadmium-induced oxidative damage in young date palm (*Phoenix dactylifera* L.). *Ecological Engineering*, 86, 202-209.
- Zulfiqar, U., Farooq, M., Hussain, S., Maqsood, M., Hussain, M., Ishfaq, M., . . . Anjum, M. Z. (2019). Lead toxicity in plants: Impacts and remediation. *Journal of environmental management*, 250, 109557.



- Zvereva, E. L., & Kozlov, M. V. (2011). Impacts of industrial polluters on bryophytes: a meta-analysis of observational studies. *Water, Air, & Soil Pollution*, 218, 573-586.
- Zwolak, A., Sarzyńska, M., Szpyrka, E., & Stawarczyk, K. (2019). Sources of soil pollution by heavy metals and their accumulation in vegetables: A review. *Water, air, & soil pollution*, 230, 1-9.



## Appendix

Appendix Table 1. Plant species list with their families and phytogeographic elements of the study area.

Families	Habit	Species	Phytogeographic elements
<b>Acanthaceae</b>	Herb	<i>Justica nilgherrensis</i>	Asiatic
<b>Amaranthaceae</b>	Herb	<i>Achyranthes aspera</i>	Tropical+ subtropical
	Herb	<i>Amaranthus graecizans L.</i>	Neotropical
	Herb	<i>Amaranthus retroflexus</i>	Cosmopolitan
	Herb	<i>Amaranthus spinosus L.</i>	Neotropical
	Herb	<i>Amaranthus viridis L.</i>	Cosmopolitan
	Herb	<i>Chenopodium album L.</i>	Cosmopolitan
	Herb	<i>Dysphania ambrosioides (L.) Mosyakin &amp; Clemants</i>	American
	Herb	<i>Salsola Kali L.</i>	Cosmopolitan
<b>Anacardiaceae</b>	Tree	<i>Mangifera indica L.</i>	Pantropical
<b>Apiaceae</b>	Herb	<i>Heracleum sphondylium L.</i>	Eurasia
	Herb	<i>Torilis japonica (Houtt.) DC</i>	Pluriregional
	Herb	<i>Trifolium resupinatum L.</i>	Mediterranean + Irano-Turanian+ Euro-Siberian
	Herb	<i>Torilis leptophylla (L.) Rchb.f</i>	Irano-Turanian
<b>Apocynaceae</b>	Shrub	<i>Calotropis procera (Aiton) Dryand.</i>	Pluriregional
<b>Araceae</b>	Herb	<i>Alocasia macrorrhizos (L.) G.Don</i>	Tropical+ subtropical
	Herb	<i>Colocasia esculenta (L.) schott</i>	Tropical
	Herb	<i>Colocasia gigantea (Blume) Hook.f.</i>	Indo-Asia
	Herb	<i>Lemna minor L.</i>	Pluriregional



	Herb	<i>Leuccocasia gigantea</i> (Blume) Schott	Indo-Asia
	Herb	<i>Pistia stratiotes</i> L.	Pantropical
<b>Asparagaceae</b>	Herb	<i>Echeandia reflexa</i> (Cav.) Rose	American
<b>Brassicaceae</b>	Herb	<i>Brassica compestris</i> L.	Irano-Turanian+Sino Japanese
	Herb	<i>Brassica oleracea</i> L.	Cosmopolitan
	Herb	<i>Coronops didymus</i> (L.) Sm.	Cosmopolitan
	Herb	<i>Goldbachia laevigata</i> (M. Bieb.) DC.	Irano-Turanian+ Mediterranean
	Herb	<i>Nastrum officinales</i> R. Br.	Irano-Turanian+Sino- Japanese
	Herb	<i>Sinapis arvensis</i> L.	Cosmopolitan
	Herb	<i>Sisimbrium irio</i>	Eurasia
<b>Cannabaceae</b>	Shrub	<i>Cannabis sativa</i> L.	Irano-Turanian
<b>Caryophyllaceae</b>	Herb	<i>Stellaria media</i> (L.) L.	Eurasia
<b>Commelinaceae</b>	Herb	<i>Commelina benghalensis</i>	Tropical
<b>Compositae</b>	Herb	<i>Ageratum conyzoides</i> L.	Neotropical
	Herb	<i>Artemisia brevifolia</i>	Irano-Turanian
	Herb	<i>Artemisia scoparia</i> Waldst. & Kitam	Irano-Turanian
	Herb	<i>Cichorium intybus</i> L.	Irano-Turanian+ Saharo- Sindian
	Herb	<i>Conyza bonariensis</i> var. <i>Leiantha</i>	Saharo-Arabian
	Herb	<i>Conyza canadensis</i> (L.) Cronq.	Pluriregional
	Herb	<i>Echinops latifolius</i>	East Asia
	Herb	<i>Eclipta alba</i> (L.) Hassk	Neotropical
	Herb	<i>Erigeron canadensis</i> L.	Pluriregional
	Herb	<i>Erigeron bonariensis</i> L.	Pantropical
	Herb	<i>Jurinea heteromalla</i>	Irano-Turanian+ Himalaya



		(D.Don) N.Garcia, Herrando	
	Herb	<i>Parthenium hysterophorus</i> L.	Irano-Turanian
	Herb	<i>Silybum marianum</i> (L.) Gaertn.	Eurasia
	Herb	<i>Sonchus asper</i>	Eurasia
	Herb	<i>Sonchus oleraceus</i>	Mediterranean
	Herb	<i>Sylibum marianum</i> (L.) Gaertn.	Irano-Turanian
	Herb	<i>Taraxacum officinale</i>	Cosmopolitan
	Herb	<i>Xanthium strumarium</i> L.	Pantropical
<b>Convolvulaceae</b>	Herb	<i>Convolvulus arvensis</i> L.	Irano-Turanian+ Sino Japanese
	Herb	<i>Ipomia purpurea</i> (L.) Roth	Pantropical
	Herb	<i>Ipomoea carnea</i> Jacq.	Tropical
<b>Cyperaceae</b>	Herb	<i>Cyperus rotundus</i> L.	Tropical+ subtropical
<b>Elatinaceae</b>	Herb	<i>Bergia capensis</i> L.	Palaeotropical
<b>Euphorbiaceae</b>	Herb	<i>Euphorbia helioscopia</i> L.	Irano-Turanian
	Shrub	<i>Ricinus communis</i> L.	Cosmopolitan+ Pantropical
	Herb	<i>Euphorbia hirta</i>	Pantropical
<b>Fabaceae</b>	Tree	<i>Acacia homalophylla</i> Medic.	Tropical
	Tree	<i>Acacia nilotica</i> (L.) Delile	Paleotropical
	Herb	<i>Cassia occidentalis</i> L.	Pluriregional
	Tree	<i>Dalbergia sisso</i> DC	Tropical
	Herb	<i>Lathyrus pseudocicera</i> Pomp.	Mediterranean
	Tree	<i>Erythrina crista-galli</i> L.	Tropical
	Herb	<i>Indigofera linifolia</i> (L.f.) Retz.	Tropical
	Herb	<i>Lathyrus aphaca</i> L.	Irano-Turanian
	Herb	<i>Medicago denticulata</i> Willd	Holoarctic



	Herb	<i>Medicago minima</i> (L.) <i>Grub.</i>	Sino-Japanese
	Herb	<i>Medicago polymorpha</i>	Mediterranean
	Tree	<i>Prosopis juliflora</i> (Sw.) DC	Tropical
	Herb	<i>Rifolium microdon</i> Hook. & <i>Arn.</i>	America
	Herb	<i>Senna occidentalis</i> (L.) Link	Neotropical
	Herb	<i>Trifolium alexandrinum</i> L.	Asiatic
	Herb	<i>Vicia sativa</i> L.	Eurasia
<b>Juncaceae</b>	Herb	<i>Juncus acuminatus</i>	American
	Herb	<i>Juncus effuses</i> L.	Pantropical
<b>Lamiaceae</b>	Herb	<i>Clinopodium umbrosum</i> (M.Bieb.) Kuntze.	Eurasia
	Herb	<i>Mentha spicata</i> L.	Cosmopolitan
<b>Malvaceae</b>	Herb	<i>Abelmoschus moschatus</i> <i>Medic.</i>	Tropical
	Herb	<i>Abutilon indicum</i>	Tropical
	Herb	<i>Malava neglecta</i> Wallr.	Irano-Turanian
	Herb	<i>Malvastrum</i> <i>coromandelianum</i> (L.) <i>Garcke</i>	Pantropical
	Herb	<i>Sida cordata</i> (Burm.f.) <i>Borss. Waalk</i>	Tropical+subtropical
<b>Marsileaceae</b>	Herb	<i>Marsilea mutica</i> Mett.	Australian
<b>Meliaceae</b>	Tree	<i>Melia azedarach</i> L.	Irano-Turanian+ Sino-Japanese
<b>Moraceae</b>	Tree	<i>Brossunatia papyrifera</i> (L.) <i>L'Hér. ex-Vent.</i>	East Asian
	Tree	<i>Ficus benghalensis</i> L.	Irano-Turanian
	Tree	<i>Ficus carica</i> L.	Mediterranean
	Tree	<i>Ficus elastica</i> Roxb.	Tropical
	Tree	<i>Ficus virens</i> Aiton.	Tropical
	Tree	<i>Ficus religiosa</i> L.	Indo-Chinese



	Tree	<i>Morus alba L.</i>	East Asian
	Tree	<i>Morus nigra L.</i>	Irano-Turanian
<b>Myrtaceae</b>	Tree	<i>Eucalyptus camaldulensis</i> <i>Labill.</i>	Australian
	Tree	<i>Eucalyptus globulus Labill.</i>	Australian
	Tree	<i>Syzygium cumini (L.) Skeels</i>	Indo-Malaysian
<b>Nyctaginaceae</b>	Herb	<i>Boerhavia procumbens</i> <i>Banks ex Roxb.</i>	Tropical
<b>Oxalidaceae</b>	Herb	<i>Oxalis corniculata L.</i>	Cosmopolitan
<b>Papaveraceae</b>	Herb	<i>Argemone mexicana L.</i>	Tropical
	Herb	<i>Fumaria indica (Hausskn.)</i> <i>Pugsley</i>	Irano-Turanian
<b>Plantaginaceae</b>	Herb	<i>Campylanthus</i> <i>ramosissimus Wight</i>	Pakistan
	Herb	<i>Veronica anagallis-</i> <i>aquatica L.</i>	Cosmopolitan
<b>Poaceae</b>	Shrub	<i>Arundo donax L.</i>	Cosmopolitan
	Herb	<i>Avena sativa L.</i>	Pluriregional
	Herb	<i>Brachiaria reptans L.</i>	Tropical
	Tree	<i>Bromus japonicus Thunb.</i>	Pluriregional
	Herb	<i>Cenchrus biflorus Roxb.</i>	Tropical
	Herb	<i>Cenchrus ciliaris L.</i>	Cosmopolitan
	Herb	<i>Cynodon dactylon (L.)</i> <i>Pers.</i>	Cosmopolitan
	Herb	<i>Cynodon radiatus Roth.</i>	Tropical+subtropical
	Herb	<i>Desmostachy bippinanta</i> <i>(L.) Stapf</i>	Paleotropical
	Herb	<i>Dicanthium annulatum</i> <i>(Forssk) Stapf</i>	Tropical+ subtropical
	Herb	<i>Imperata cylindrical (L.)</i> <i>Raeusch.</i>	Pantropical
	Herb	<i>Koeleria macarantha</i> <i>(Ledeb.) Schult.</i>	Pluriregional



	Herb	<i>Paspalum paspalodes</i> (Michx.)	Tropical+ subtropical
	Herb	<i>Phragmites karka</i> (Retz.) <i>Trin. ex Steud.</i>	Cosmopolitan
	Shrub	<i>Saccharum bengalensis</i> <i>Retz</i>	Pluriregional
	Shrub	<i>Saccharum spontaneum</i> L.	Pluriregional
	Herb	<i>Setaria pumila</i> (Poir.) <i>Roem. &amp; Schult</i>	Pluriregional
	Herb	<i>Sorghum halepense</i> (L.) <i>Pers</i>	Cultivated
	Herb	<i>Triticum aestivum</i>	Cosmopolitan
<b>Polygonaceae</b>	Herb	<i>Persiaris glabra</i> L.	Pantropical
	Herb	<i>Rumex dentatus</i> L.	Irano-Turanian+ Sino-Japanese
	Herb	<i>Rumex nepalensis</i> Spreng.	Mediterranean
<b>Pontederiaceae</b>	Herb	<i>Eichhornia crassipes</i> (Mart.) Solms-Laub	Neotropical
<b>Primulaceae</b>	Herb	<i>Anagallis arvensis</i> L.	Cosmopolitan
<b>Pteridaceae</b>	Herb	<i>Adiantum capillus-veneris</i> L.	Subcosmopolitan
<b>Ranunculaceae</b>	Herb	<i>Ranunculus muricatus</i> L.	Irano-Turanian+ Sino-Japanese
<b>Rhamnaceae</b>	Tree	<i>Ziziphus jujuba</i> Mill.	Sino-Janpanese
	Tree	<i>Ziziphus mauritania</i>	Paleotropical
	Tree	<i>Ziziphus nummularia</i> (Burm.f.) Wight & Arn.	Paleotropical
<b>Rosaceae</b>	Herb	<i>Geum urbanum</i> L.	Irano-Turanian+ Euro-Siberian
<b>Rubiaceae</b>	Herb	<i>Galium aparine</i> L.	Holoarctic
<b>Salicaceae</b>	Tree	<i>Populus alba</i> L.	Cosmopolitan
	Tree	<i>Populus nigra</i> L.	Sino-Janpanese
	Tree	<i>Salix tetrasperma</i> Roxb.	Sub-Himalyan



<b>Scrophulariaceae</b>	Herb	<i>Verbena bonariensis L.</i>	American
	Herb	<i>Verbascum songaricum Schrenk</i>	Western Himalaya
	Herb	<i>Verbascum thapsus L.</i>	Eurasia
<b>Solanaceae</b>	Herb	<i>Datura alba L.</i>	Pluriregional
	Herb	<i>Datura innoxia</i>	Tropical+ subtropical
	Herb	<i>Solanum lycopersicum L.</i>	American
	Herb	<i>Solanum nigrum</i>	Cosmopolitan
	Herb	<i>Withania somnifera (L.) Dunal</i>	Pluriregional
<b>Tamaricaceae</b>	Tree	<i>Tamarix dioica Roxb. Ex Roth</i>	Eurasia
<b>Thymelaeaceae</b>	Herb	<i>Daphane macronata Royle</i>	Irano-Turanian
<b>Typhaceae</b>	Herb	<i>Typha angustifolia L.</i>	Holoarctic
<b>Verbenaceae</b>	Shrub	<i>Lantana camara L.</i>	Sino-Japanese
	Herb	<i>Verbena officinale</i>	Paleotropical
	Herb	<i>Verb supina L.</i>	Saharo-Sindian + Irano-Turanian



Appendix table 2 Calculated CCME-WQI along with water status of the Aik-Stream

S31	upstream	74.745	Fair
S32	upstream	73.347	Fair
S33	upstream	65.428	Fair
S34	upstream	66.825	Fair
S35	upstream	61.4805	Fair
S36	upstream	68.836	Fair
S37	upstream	67.813	Fair
S38	upstream	67.223	Fair
S39	upstream	74.745	Fair
S40	upstream	73.347	Fair
S41	upstream	65.428	Fair
S42	upstream	66.825	Fair
S43	upstream	61.4805	Marginal
S44	upstream	68.836	Fair
S45	upstream	64.813	Marginal
S46	upstream	64.223	Marginal
S47	upstream	65.764	Fair
S48	upstream	62.223	Marginal
S49	upstream	68.347	Fair
S50	upstream	60.428	Marginal
S51	upstream	61.825	Marginal
S52	upstream	38.91	Poor
S53	midstream	38.801	Poor
S54	midstream	39.493	Poor
S55	midstream	34.152	Poor
S56	midstream	38.517	Poor
S57	midstream	35.206	Poor
S58	midstream	36.734	Poor
S59	midstream	34.693	Poor
S60	midstream	34.845	Poor
S61	midstream	35.882	Poor
S62	midstream	36.266	Poor

Sampling sites	Location	WQI	Status
S1	upstream	85.996	Good
S2	upstream	86.344	Good
S3	upstream	77.174	Fair
S4	upstream	81.253	Good
S5	upstream	76.745	Fair
S6	upstream	75.347	Fair
S7	upstream	67.428	Fair
S8	upstream	68.825	Fair
S9	upstream	63.4805	Marginal
S10	upstream	70.836	Fair
S11	upstream	69.813	Fair
S12	upstream	69.223	Fair
S13	upstream	70.764	Fair
S14	upstream	69.223	Fair
S15	upstream	64.586	Fair
S16	upstream	72.382	Fair
S17	upstream	69.074	Fair
S18	upstream	78.745	Fair
S19	upstream	77.347	Fair
S20	upstream	69.428	Fair
S21	upstream	70.825	Fair
S22	upstream	65.4805	Fair
S23	upstream	72.836	Fair
S24	upstream	71.813	Fair
S25	upstream	71.223	Fair
S26	upstream	72.764	Fair
S27	upstream	72.223	Fair
S28	upstream	67.586	Fair
S29	upstream	75.382	Fair
S30	upstream	67.074	Fair



<b>S97</b>	midstream	31.948	<b>Poor</b>
<b>S98</b>	midstream	30.942	<b>Poor</b>
<b>S99</b>	midstream	35.307	<b>Poor</b>
<b>S100</b>	midstream	31.996	<b>Poor</b>
<b>S101</b>	midstream	33.524	<b>Poor</b>
<b>S102</b>	midstream	31.483	<b>Poor</b>
<b>S103</b>	midstream	31.635	<b>Poor</b>
<b>S104</b>	midstream	32.672	<b>Poor</b>
<b>S105</b>	midstream	33.056	<b>Poor</b>
<b>S106</b>	midstream	32.546	<b>Poor</b>
<b>S107</b>	midstream	31.948	<b>Poor</b>
<b>S108</b>	midstream	31.632	<b>Poor</b>
<b>S109</b>	midstream	35.997	<b>Poor</b>
<b>S110</b>	midstream	32.686	<b>Poor</b>
<b>S111</b>	midstream	34.214	<b>Poor</b>
<b>S112</b>	midstream	32.173	<b>Poor</b>
<b>S113</b>	midstream	32.325	<b>Poor</b>
<b>S114</b>	midstream	33.362	<b>Poor</b>
<b>S115</b>	midstream	33.746	<b>Poor</b>
<b>S116</b>	midstream	33.236	<b>Poor</b>
<b>S117</b>	midstream	35.997	<b>Poor</b>
<b>S118</b>	downstream	46.199	<b>Marginal</b>
<b>S119</b>	downstream	46.409	<b>Marginal</b>
<b>S120</b>	downstream	45.812	<b>Marginal</b>
<b>S121</b>	downstream	48.409	<b>Marginal</b>
<b>S122</b>	downstream	47.688	<b>Marginal</b>
<b>S123</b>	downstream	49.717	<b>Marginal</b>
<b>S124</b>	downstream	48.955	<b>Marginal</b>
<b>S125</b>	downstream	48.261	<b>Marginal</b>
<b>S126</b>	downstream	49.005	<b>Marginal</b>
<b>S127</b>	downstream	48.345	<b>Marginal</b>
<b>S128</b>	downstream	47.301	<b>Marginal</b>
<b>S129</b>	downstream	48.949	<b>Marginal</b>
<b>S130</b>	downstream	49.159	<b>Marginal</b>

<b>S63</b>	midstream	35.756	<b>Poor</b>
<b>S64</b>	midstream	35.158	<b>Poor</b>
<b>S65</b>	midstream	35.329	<b>Poor</b>
<b>S66</b>	midstream	35.548	<b>Poor</b>
<b>S67</b>	midstream	32.546	<b>Poor</b>
<b>S68</b>	midstream	29.837	<b>Poor</b>
<b>S69</b>	midstream	37.733	<b>Poor</b>
<b>S70</b>	midstream	34.582	<b>Poor</b>
<b>S71</b>	midstream	35.838	<b>Poor</b>
<b>S72</b>	midstream	36.082	<b>Poor</b>
<b>S73</b>	midstream	36.112	<b>Poor</b>
<b>S74</b>	midstream	36.801	<b>Poor</b>
<b>S75</b>	midstream	37.493	<b>Poor</b>
<b>S76</b>	midstream	32.152	<b>Poor</b>
<b>S77</b>	midstream	36.517	<b>Poor</b>
<b>S78</b>	midstream	33.206	<b>Poor</b>
<b>S79</b>	midstream	34.734	<b>Poor</b>
<b>S80</b>	midstream	32.693	<b>Poor</b>
<b>S81</b>	midstream	32.845	<b>Poor</b>
<b>S82</b>	midstream	33.882	<b>Poor</b>
<b>S83</b>	midstream	34.266	<b>Poor</b>
<b>S84</b>	midstream	33.756	<b>Poor</b>
<b>S85</b>	midstream	33.158	<b>Poor</b>
<b>S86</b>	midstream	35.801	<b>Poor</b>
<b>S87</b>	midstream	36.493	<b>Poor</b>
<b>S88</b>	midstream	31.152	<b>Poor</b>
<b>S89</b>	midstream	35.517	<b>Poor</b>
<b>S90</b>	midstream	32.206	<b>Poor</b>
<b>S91</b>	midstream	33.734	<b>Poor</b>
<b>S92</b>	midstream	31.693	<b>Poor</b>
<b>S93</b>	midstream	31.845	<b>Poor</b>
<b>S94</b>	midstream	32.882	<b>Poor</b>
<b>S95</b>	midstream	32.266	<b>Poor</b>
<b>S96</b>	midstream	32.546	<b>Poor</b>



<b>S141</b>	downstream	50.898	<b>Marginal</b>
<b>S142</b>	downstream	52.927	<b>Marginal</b>
<b>S143</b>	downstream	52.165	<b>Marginal</b>
<b>S144</b>	downstream	51.471	<b>Marginal</b>
<b>S145</b>	downstream	52.215	<b>Marginal</b>
<b>S146</b>	downstream	51.555	<b>Marginal</b>
<b>S147</b>	downstream	50.511	<b>Marginal</b>
<b>S148</b>	downstream	49.409	<b>Marginal</b>
<b>S149</b>	downstream	49.619	<b>Marginal</b>
<b>S150</b>	downstream	49.022	<b>Marginal</b>

<b>S131</b>	downstream	48.562	<b>Marginal</b>
<b>S132</b>	downstream	51.159	<b>Marginal</b>
<b>S133</b>	downstream	50.438	<b>Marginal</b>
<b>S134</b>	downstream	52.467	<b>Marginal</b>
<b>S135</b>	downstream	51.705	<b>Marginal</b>
<b>S136</b>	downstream	51.011	<b>Marginal</b>
<b>S137</b>	downstream	51.755	<b>Marginal</b>
<b>S138</b>	downstream	51.095	<b>Marginal</b>
<b>S139</b>	downstream	50.051	<b>Marginal</b>
<b>S140</b>	downstream	51.619	<b>Marginal</b>



Appendix table 3 Contamination Factor (CF) and Degree of Contamination (DC) of heavy metals at the soil around the Aik-Stream.

S	3	4	0	8	0	3	9	5	0	2	1	Contamination Factor (CF)								DC
S39	0.408	0.395	0.287	0.172	0.306	0.039	0.010	0.022	1.639	S	S	Cu	Zn	Cr	Pb	Cd	Ni	As	Hg	
S40	0.285	0.447	0.226	0.109	0.943	0.094	0.011	0.041	2.156	S1		0.149	0.223	0.254	0.144	0.803	0.078	0.007	0.008	1.665
S41	0.257	0.428	0.305	0.087	0.803	0.062	0.014	0.055	2.011	S2		0.148	0.203	0.226	0.179	0.643	0.092	0.005	0.012	1.509
S42	0.257	0.487	0.275	0.096	1.089	0.056	0.015	0.018	2.293	S3		0.287	0.443	0.385	0.120	0.545	0.085	0.009	0.017	1.892
S43	0.308	0.412	0.357	0.049	0.780	0.079	0.016	0.034	2.036	S4		0.237	0.473	0.216	0.143	0.890	0.066	0.007	0.021	2.053
S44	0.365	0.434	0.337	0.049	0.306	0.056	0.020	0.038	1.605	S5		0.182	0.433	0.225	0.131	0.803	0.076	0.009	0.021	1.879
S45	0.393	0.393	0.366	0.072	0.803	0.044	0.019	0.034	2.124	S6		0.237	0.374	0.346	0.109	1.053	0.076	0.010	0.022	2.226
S46	0.307	0.348	0.307	0.101	0.056	0.038	0.022	0.024	1.203	S7		0.149	0.520	0.330	0.099	0.068	0.060	0.011	0.012	1.248
S47	0.335	0.488	0.375	0.149	0.068	0.084	0.023	0.045	1.566	S8		0.288	0.259	0.279	0.052	0.080	0.089	0.011	0.041	1.100
S48	0.279	0.547	0.249	0.186	0.193	0.052	0.020	0.031	1.558	S9		0.233	0.299	0.288	0.099	0.939	0.096	0.014	0.055	2.023
S49	0.326	0.538	0.304	0.158	0.818	0.057	0.019	0.024	2.243	S10		0.312	0.233	0.270	0.075	0.264	0.058	0.015	0.018	1.244
S50	0.326	0.498	0.313	0.157	0.803	0.041	0.022	0.045	2.205	S11		0.284	0.214	0.244	0.156	0.680	0.047	0.016	0.034	1.675
S51	0.346	0.500	0.317	0.158	1.431	0.047	0.020	0.039	2.858	S12		0.395	0.492	0.306	0.125	0.443	0.044	0.020	0.038	1.863
S52	0.376	0.467	0.469	1.022	0.678	1.221	0.089	0.094	4.414	S13		0.458	0.437	0.320	0.130	0.138	0.067	0.019	0.034	1.603
S53	0.376	0.666	0.879	1.064	1.951	1.204	0.096	0.095	6.331	S14		0.486	0.377	0.289	0.031	0.105	0.039	0.022	0.024	1.373
S54	0.404	0.628	0.668	1.063	1.431	1.261	0.089	0.074	5.618	S15		0.429	0.377	0.336	0.077	2.065	0.018	0.023	0.045	3.371
S55	0.583	0.475	0.718	1.064	2.938	1.234	0.115	0.928	8.054	S16		0.379	0.548	0.339	0.107	1.651	0.028	0.020	0.031	3.103
S56	0.387	0.455	0.656	1.050	3.306	1.240	0.157	0.125	7.377	S17		0.323	0.468	0.308	0.190	0.810	0.047	0.007	0.008	2.161
S57	0.593	0.596	0.597	1.028	3.315	1.229	0.111	0.181	7.649	S18		0.240	0.487	0.289	0.180	1.193	0.010	0.005	0.012	2.415
S58	0.655	0.687	0.424	1.014	1.983	1.145	0.138	0.097	6.143	S19		0.405	0.559	0.317	0.103	0.771	0.015	0.009	0.017	2.196
S59	0.722	0.627	0.384	1.013	2.820	1.173	0.198	0.356	7.294	S20		0.405	0.579	0.294	0.174	2.053	0.045	0.007	0.021	3.576
S60	0.798	0.608	0.524	1.128	2.051	1.173	0.164	0.446	6.892	S21		0.348	0.435	0.306	0.137	1.781	0.016	0.009	0.021	3.054
S61	0.903	0.667	0.507	1.060	2.434	1.260	0.172	0.385	7.389	S22		0.327	0.409	0.280	0.090	0.105	0.032	0.010	0.022	1.275
S62	0.903	0.787	0.516	1.172	2.938	1.236	0.195	0.172	7.919	S23		0.326	0.350	0.246	0.102	0.776	0.039	0.011	0.012	1.860
S63	0.655	0.707	0.488	1.174	4.314	1.520	0.164	0.542	9.563	S24		0.327	0.488	0.265	0.074	1.068	0.038	0.011	0.041	2.313
S64	0.648	0.820	0.493	1.164	3.574	1.202	0.184	0.342	8.427	S25		0.302	0.473	0.313	0.059	0.765	0.102	0.014	0.055	2.082
S65	0.606	0.815	0.514	1.003	2.928	1.179	0.197	0.245	7.486	S26		0.357	0.433	0.354	0.094	0.195	0.102	0.015	0.018	1.568
S66	0.704	0.914	0.382	1.031	2.801	1.144	0.175	0.164	7.316	S27		0.329	0.489	0.323	0.080	0.926	0.046	0.016	0.034	2.244
S67	0.657	0.835	0.430	1.052	2.960	0.972	0.183	0.195	7.284	S28		0.273	0.435	0.369	0.161	1.051	0.081	0.020	0.038	2.428
S68	0.879	0.835	0.509	0.952	3.943	1.173	0.134	0.172	8.597	S29		0.329	0.468	0.350	0.137	0.318	0.052	0.019	0.034	1.707
S69	0.735	0.915	0.558	1.029	2.793	1.153	0.164	0.194	7.541	S30		0.329	0.488	0.378	0.113	0.818	0.068	0.022	0.024	2.240
S70	0.895	1.230	1.618	1.084	2.343	1.178	0.167	0.416	8.930	S31		0.304	0.513	0.326	0.095	1.234	0.010	0.023	0.045	2.550
S71	0.618	1.284	1.596	1.083	2.808	1.299	0.089	0.368	9.145	S32		0.415	0.512	0.246	0.083	0.710	0.067	0.020	0.031	2.086
S72	0.618	1.246	0.815	1.090	3.615	1.167	0.123	0.782	9.456	S33		0.248	0.395	0.235	0.071	1.155	0.068	0.007	0.008	2.187
S73	1.130	1.113	1.236	1.133	4.196	1.209	0.020	0.039	10.077	S34		0.230	0.489	0.222	0.106	0.289	0.092	0.005	0.012	1.444
S74	1.095	0.712	0.936	1.079	2.931	1.404	0.089	0.094	8.340	S35		0.312	0.569	0.231	0.070	0.193	0.076	0.009	0.017	1.477
S75	1.187	0.950	1.144	1.079	3.101	1.373	0.096	0.095	9.026	S36		0.340	0.490	0.242	0.127	0.063	0.044	0.007	0.021	1.333
S76	1.193	0.816	0.860	1.100	2.724	1.344	0.089	0.074	8.199	S37		0.185	0.553	0.208	0.137	1.156	0.038	0.009	0.021	2.308



<b>S115</b>	1.809	1.113	2.448	1.039	2.140	0.371	0.089	0.368	9.377	<b>S77</b>	1.248	0.999	1.760	1.113	1.568	1.382	0.115	0.928	9.112
<b>S116</b>	1.269	0.892	1.443	0.992	1.815	0.313	0.123	0.782	7.629	<b>S78</b>	1.109	0.915	1.058	1.078	3.026	1.333	0.157	0.125	8.800
<b>S117</b>	0.984	1.270	1.401	1.129	2.064	0.356	0.023	0.251	7.478	<b>S79</b>	1.041	1.102	1.252	1.090	2.300	1.230	0.111	0.181	8.307
<b>S118</b>	0.871	0.843	1.006	0.471	0.106	0.298	0.037	0.132	3.765	<b>S80</b>	1.040	1.023	1.752	1.096	2.808	1.271	0.138	0.097	9.225
<b>S119</b>	0.872	0.849	1.064	0.490	0.068	0.263	0.026	0.216	3.849	<b>S81</b>	1.067	1.084	1.554	1.086	1.651	1.202	0.198	0.356	8.199
<b>S120</b>	0.926	0.507	0.885	0.549	0.081	0.239	0.042	0.147	3.376	<b>S82</b>	1.075	1.099	1.061	1.116	3.060	1.151	0.164	0.446	9.172
<b>S121</b>	1.073	1.021	0.827	0.604	1.215	0.336	0.016	0.334	5.428	<b>S83</b>	1.048	1.149	1.561	1.102	4.453	1.219	0.172	0.385	11.087
<b>S122</b>	1.158	0.571	0.737	0.596	0.053	0.267	0.024	0.146	3.551	<b>S84</b>	1.076	1.113	1.662	1.084	4.195	1.204	0.195	0.172	10.701
<b>S123</b>	1.156	0.593	0.419	0.631	0.565	0.212	0.031	0.092	3.700	<b>S85</b>	0.932	1.087	1.774	1.117	5.845	0.989	0.164	0.542	12.452
<b>S124</b>	0.796	0.413	0.729	0.480	0.068	0.116	0.023	0.125	2.750	<b>S86</b>	1.014	0.966	1.545	1.087	4.018	1.286	0.184	0.342	10.442
<b>S125</b>	0.824	0.553	0.338	0.556	0.046	0.150	0.028	0.229	2.723	<b>S87</b>	1.294	0.943	1.473	1.099	2.818	1.144	0.197	0.245	9.213
<b>S126</b>	0.767	0.655	0.726	0.502	0.349	0.148	0.056	0.122	3.325	<b>S88</b>	0.763	1.229	1.007	1.086	6.084	1.196	0.175	0.164	11.704
<b>S127</b>	0.882	0.453	0.559	0.437	0.098	0.235	0.052	0.224	2.940	<b>S89</b>	0.795	1.215	1.018	1.077	3.314	1.293	0.183	0.195	9.090
<b>S128</b>	0.882	0.935	0.643	0.396	0.855	0.229	0.023	0.251	4.214	<b>S90</b>	1.188	1.206	1.295	1.082	4.426	1.293	0.134	0.172	10.795
<b>S129</b>	0.939	0.957	0.229	0.397	1.205	0.218	0.037	0.132	4.116	<b>S91</b>	1.228	1.189	1.216	1.090	3.176	1.144	0.164	0.194	9.402
<b>S130</b>	0.610	0.813	0.346	0.246	0.056	0.114	0.026	0.216	2.427	<b>S92</b>	1.258	1.168	1.354	1.058	4.051	1.245	0.167	0.416	10.718
<b>S131</b>	0.609	1.025	0.427	0.243	0.059	0.124	0.042	0.147	2.677	<b>S93</b>	1.201	1.085	1.346	1.090	4.018	1.218	0.089	0.368	10.415
<b>S132</b>	0.554	1.012	0.332	0.257	1.178	0.133	0.016	0.334	3.816	<b>S94</b>	1.206	1.310	1.354	1.015	4.428	1.240	0.123	0.782	11.457
<b>S133</b>	0.512	0.893	0.230	0.169	1.089	0.165	0.024	0.146	3.228	<b>S95</b>	1.207	1.290	1.449	1.004	2.818	1.246	0.020	0.039	9.073
<b>S134</b>	0.485	0.891	0.520	0.331	0.955	0.301	0.031	0.092	3.606	<b>S96</b>	0.938	1.344	1.223	1.017	5.401	1.265	0.089	0.094	11.370
<b>S135</b>	0.629	0.740	0.406	0.364	0.268	0.296	0.023	0.125	2.852	<b>S97</b>	1.238	1.355	1.297	1.117	5.439	1.232	0.096	0.095	11.868
<b>S136</b>	0.709	0.653	0.440	0.392	0.928	0.346	0.028	0.229	3.725	<b>S98</b>	1.237	1.257	1.295	1.105	4.064	1.262	0.089	0.074	10.383
<b>S137</b>	0.346	0.657	0.420	0.321	0.405	0.343	0.056	0.122	2.671	<b>S99</b>	1.315	1.272	1.194	1.090	2.806	1.241	0.115	0.928	9.961
<b>S138</b>	0.567	0.782	0.319	0.219	1.231	0.284	0.052	0.224	3.680	<b>S100</b>	1.209	1.009	2.064	1.018	5.406	1.382	0.157	0.125	12.370
<b>S139</b>	0.592	0.681	0.221	0.174	0.026	0.225	0.023	0.251	2.195	<b>S101</b>	1.211	1.115	2.154	1.015	2.945	1.350	0.111	0.181	10.081
<b>S140</b>	0.537	0.621	0.710	0.162	0.943	0.204	0.037	0.132	3.346	<b>S102</b>	1.187	1.088	2.357	1.003	1.568	1.304	0.138	0.097	8.743
<b>S141</b>	0.435	0.813	0.510	0.267	0.080	0.161	0.026	0.216	2.509	<b>S103</b>	1.174	1.548	2.661	1.016	3.103	1.156	0.198	0.356	11.211
<b>S142</b>	0.574	0.830	0.326	0.225	1.105	0.121	0.042	0.147	3.370	<b>S104</b>	1.174	1.453	2.443	1.090	5.444	1.502	0.164	0.446	13.717
<b>S143</b>	0.657	0.887	0.313	0.377	0.431	0.118	0.016	0.334	3.135	<b>S105</b>	1.730	1.476	2.432	1.114	4.433	1.276	0.172	0.385	13.017
<b>S144</b>	0.574	0.679	0.406	0.149	1.193	0.175	0.024	0.146	3.346	<b>S106</b>	1.713	1.609	2.055	1.052	3.318	1.254	0.195	0.172	11.367
<b>S145</b>	0.421	0.511	0.195	0.205	0.146	0.094	0.031	0.092	1.696	<b>S107</b>	1.435	1.630	2.116	0.964	1.695	1.276	0.164	0.542	9.823
<b>S146</b>	0.616	0.533	0.372	0.375	0.818	0.235	0.023	0.125	3.096	<b>S108</b>	1.436	1.211	2.238	1.019	1.519	1.337	0.184	0.342	9.286
<b>S147</b>	0.426	0.568	0.381	0.302	1.228	0.322	0.028	0.229	3.483	<b>S109</b>	1.543	1.304	2.438	1.151	1.276	1.927	0.197	0.245	10.082
<b>S148</b>	0.416	0.447	0.226	0.291	0.089	0.283	0.056	0.122	1.930	<b>S110</b>	1.460	1.303	1.427	0.950	1.443	1.328	0.175	0.164	8.250
<b>S149</b>	0.304	0.399	0.316	0.209	0.981	0.158	0.052	0.224	2.644	<b>S111</b>	1.212	1.508	2.415	1.090	2.065	0.378	0.183	0.195	9.047
<b>S150</b>	0.326	0.613	0.251	0.138	0.564	0.161	0.056	0.122	2.231	<b>S112</b>	1.239	1.025	2.553	1.095	1.280	0.492	0.134	0.172	7.991
										<b>S113</b>	1.241	0.976	2.547	1.090	1.295	0.610	0.164	0.194	8.118
										<b>S114</b>	1.684	1.209	1.519	1.102	1.318	0.550	0.167	0.416	7.964



Appendix table 4 Potential Ecological Risk (PER) and Potential Ecological Risk Index (PERI) calculation of the soil around the Aik Stream

S37	2.08	0.553	1.604	47.22	3.875	0.415	0.05	0.08	55.877		Potential Ecological Risk (PER)								
S38	4.585	0.395	2.212	13.248	4.663	0.42	0.05	0.084	25.657	S. S	Cu	Zn	Cr	Cd	Pb	Ni	As	Hg	PERI
S39	3.95	0.512	1.988	8.2572	2.8872	0.61	0.05	0.0456	18.3	S1	1.675	0.223	1.956	32.4648	3.4298	0.85	0.04	0.0304	40.669
S40	3.205	0.447	1.742	39.582	4.197	1.03	0.06	0.156	50.419	S2	1.67	0.203	1.742	26.2572	4.3572	1.005	0.03	0.0456	35.31
S41	2.895	0.428	2.348	34.6152	4.3552	0.68	0.07	0.2096	45.601	S3	3.225	0.443	2.964	22.5876	3.3376	0.935	0.04	0.0648	33.597
S42	2.895	0.487	2.116	44.3856	2.8606	0.61	0.08	0.0688	53.503	S4	2.67	0.473	1.66	36.57	3.995	0.725	0.04	0.08	46.213
S43	3.47	0.412	2.744	32.7552	2.5852	0.87	0.08	0.1296	43.046	S5	2.045	0.433	1.732	33.06	3.745	0.83	0.05	0.08	41.975
S44	4.11	0.434	2.594	13.9776	2.7676	0.61	0.1	0.1448	24.738	S6	2.67	0.374	2.662	43.098	3.323	0.825	0.05	0.084	53.086
S45	4.42	0.393	2.82	33.6552	3.0852	0.48	0.1	0.1296	45.083	S7	1.68	0.52	2.538	3.2472	2.6622	0.65	0.05	0.0456	11.393
S46	3.45	0.348	2.36	3.3444	3.2344	0.415	0.11	0.0912	13.353	S8	3.24	0.259	2.148	5.082	2.987	0.97	0.06	0.156	14.902
S47	3.765	0.488	2.886	4.7544	5.2194	0.92	0.11	0.1712	18.314	S9	2.615	0.299	2.218	40.0752	4.6302	1.045	0.07	0.2096	51.162
S48	3.14	0.547	1.914	9.1308	5.3858	0.57	0.1	0.1184	20.906	S10	3.51	0.233	2.074	11.3856	2.4106	0.64	0.08	0.0688	20.402
S49	3.67	0.538	2.338	33.7944	4.4444	0.62	0.1	0.0912	45.596	S11	3.195	0.214	1.88	28.7652	4.8652	0.515	0.08	0.1296	39.644
S50	3.675	0.498	2.412	34.1544	5.3944	0.445	0.11	0.1712	46.86	S12	4.445	0.492	2.354	19.4376	4.3976	0.48	0.1	0.1448	31.851
S51	3.89	0.5	2.44	59.0256	5.1406	0.515	0.1	0.1488	71.76	S13	5.155	0.437	2.46	7.0452	4.3252	0.73	0.1	0.1296	20.382
S52	4.23	0.467	3.604	31.3908	26.0108	13.355	0.44	0.3584	79.856	S14	5.465	0.377	2.22	5.2944	1.7494	0.425	0.11	0.0912	15.732
S53	4.23	0.666	6.758	82.3992	26.9442	13.17	0.48	0.3616	135.009	S15	4.83	0.377	2.586	84.6444	3.6894	0.2	0.11	0.1712	96.608
S54	4.54	0.628	5.142	60.6192	25.9592	13.79	0.45	0.2816	111.41	S16	4.265	0.548	2.606	67.4808	3.6958	0.305	0.1	0.1184	79.119
S55	6.56	0.475	5.52	159.9324	65.0374	13.495	0.57	3.5352	255.125	S17	3.64	0.468	2.37	32.7648	4.3948	0.515	0.04	0.0304	44.223
S56	4.36	0.455	5.048	137.952	28.022	13.565	0.78	0.476	190.658	S18	2.705	0.487	2.22	48.2472	4.3622	0.11	0.03	0.0456	58.207
S57	6.67	0.596	4.594	140.8752	30.1102	13.445	0.55	0.6896	197.53	S19	4.56	0.559	2.436	31.6176	2.9626	0.165	0.04	0.0648	42.405
S58	7.37	0.687	3.266	83.7252	25.9802	12.525	0.69	0.3696	134.613	S20	4.55	0.579	2.258	83.07	4.65	0.49	0.04	0.08	95.717
S59	8.12	0.627	2.956	129.072	37.807	12.835	0.99	1.356	193.763	S21	3.915	0.435	2.358	72.21	3.87	0.175	0.05	0.08	83.093
S60	8.98	0.608	4.03	102.4404	44.3504	12.83	0.82	1.6992	175.758	S22	3.675	0.409	2.156	5.208	2.923	0.35	0.05	0.084	14.855
S61	10.165	0.667	3.898	114.9468	40.1318	13.785	0.86	1.4664	185.92	S23	3.665	0.35	1.89	31.5972	2.7072	0.425	0.05	0.0456	40.73
S62	10.165	0.787	3.972	125.3724	32.7774	13.515	0.98	0.6552	188.224	S24	3.675	0.488	2.038	44.562	3.452	0.42	0.06	0.156	54.851
S63	7.365	0.707	3.75	197.3376	49.7176	16.63	0.82	2.0648	278.392	S25	3.39	0.473	2.408	33.1152	3.7602	1.115	0.07	0.2096	44.541
S64	7.295	0.82	3.796	158.5884	40.3734	13.15	0.92	1.3032	226.246	S26	4.015	0.433	2.726	8.6256	2.8156	1.115	0.08	0.0688	19.879
S65	6.815	0.815	3.954	128.2932	32.5132	12.89	0.99	0.9336	187.204	S27	3.7	0.489	2.488	38.6052	3.2602	0.505	0.08	0.1296	49.257
S66	7.92	0.914	2.942	119.5476	29.4076	12.51	0.88	0.6248	174.746	S28	3.07	0.435	2.836	43.7976	5.1526	0.89	0.1	0.1448	56.426
S67	7.39	0.835	3.306	127.3284	31.2834	10.63	0.92	0.7432	182.436	S29	3.7	0.468	2.69	14.2452	4.4652	0.57	0.1	0.1296	26.368
S68	9.885	0.835	3.918	165.5724	28.0974	12.83	0.67	0.6552	222.463	S30	3.7	0.488	2.912	33.7944	3.4844	0.74	0.11	0.0912	45.32
S69	8.265	0.915	4.296	120.5604	30.7354	12.615	0.82	0.7392	178.946	S31	3.425	0.513	2.512	51.4044	4.0744	0.105	0.11	0.1712	62.315
S70	10.07	1.23	12.444	112.7076	42.0576	12.88	0.83	1.5848	193.804	S32	4.665	0.512	1.894	29.8308	3.1908	0.735	0.1	0.1184	41.046
S71	6.96	1.284	12.278	129.1092	39.8242	14.205	0.45	1.4016	205.512	S33	2.795	0.395	1.81	46.5648	1.8848	0.74	0.04	0.0304	54.26
S72	6.95	1.246	6.272	180.3504	58.9104	12.765	0.61	2.9792	270.083	S34	2.585	0.489	1.706	12.0972	2.7922	1.005	0.03	0.0456	20.75
S73	12.71	1.113	9.504	169.6356	25.8706	13.23	0.1	0.1488	232.312	S35	3.515	0.569	1.78	8.4876	2.2626	0.83	0.04	0.0648	17.549
S74	12.315	0.712	7.204	121.5408	27.2258	15.36	0.44	0.3584	185.156	S36	3.825	0.49	1.866	3.45	3.65	0.475	0.04	0.08	13.876



S114	18.945	1.209	11.682	71.7276	42.4276	6.015	0.83	1.5848	154.421	S75	13.36	0.95	8.802	128.3892	27.2692	15.015	0.48	0.3616	194.627
S115	20.35	1.113	18.83	102.4092	38.9042	4.06	0.45	1.4016	187.518	S76	13.425	0.816	6.612	112.3392	26.7492	14.695	0.45	0.2816	175.368
S116	14.27	0.892	11.096	108.3504	56.8404	3.425	0.61	2.9792	198.463	S77	14.045	0.999	13.536	105.1224	66.0774	15.11	0.57	3.5352	218.995
S117	11.07	1.27	10.78	94.032	35.457	3.895	0.12	0.956	157.58	S78	12.475	0.915	8.136	126.762	28.622	14.575	0.78	0.476	192.741
S118	9.795	0.843	7.738	10.2984	16.0534	3.265	0.19	0.5032	48.686	S79	11.705	1.102	9.632	100.2852	31.4402	13.46	0.55	0.6896	168.864
S119	9.81	0.849	8.186	12.5784	20.2984	2.88	0.13	0.8232	55.555	S80	11.7	1.023	13.478	116.7252	27.7252	13.9	0.69	0.3696	185.611
S120	10.42	0.507	6.81	9.96	18.38	2.61	0.21	0.56	49.457	S81	12.005	1.084	11.952	82.332	39.357	13.14	0.99	1.356	162.216
S121	12.075	1.021	6.364	63.864	28.104	3.68	0.08	1.272	116.46	S82	12.1	1.099	8.162	142.7904	44.1104	12.585	0.82	1.6992	223.366
S122	13.025	0.571	5.672	8.772	19.332	2.92	0.12	0.556	50.968	S83	11.785	1.149	12.01	195.7068	41.0068	13.33	0.86	1.4664	277.314
S123	13.01	0.593	3.226	26.7948	17.6148	2.32	0.16	0.3504	64.069	S84	12.105	1.113	12.788	175.6524	30.8874	13.17	0.98	0.6552	247.351
S124	8.955	0.413	5.612	8.412	15.902	1.265	0.12	0.476	41.155	S85	10.48	1.087	13.65	258.5676	48.5226	10.825	0.82	2.0648	346.017
S125	9.265	0.553	2.596	12.324	22.279	1.64	0.14	0.872	49.669	S86	11.41	0.966	11.884	176.3484	38.7434	14.065	0.92	1.3032	255.64
S126	8.63	0.655	5.588	19.5276	16.2426	1.62	0.28	0.4648	53.008	S87	14.555	0.943	11.332	123.9132	34.5682	12.515	0.99	0.9336	199.75
S127	9.92	0.453	4.304	14.1432	19.5332	2.565	0.26	0.8536	52.032	S88	8.58	1.229	7.742	250.8576	30.5826	13.075	0.88	0.6248	313.571
S128	9.925	0.935	4.942	45.672	19.887	2.51	0.12	0.956	84.947	S89	8.94	1.215	7.83	141.4584	31.8034	14.145	0.92	0.7432	207.055
S129	10.57	0.957	1.764	54.2484	14.4784	2.39	0.19	0.5032	85.101	S90	13.365	1.206	9.96	184.9224	30.8524	14.14	0.67	0.6552	255.771
S130	6.86	0.813	2.664	12.1284	15.1034	1.245	0.13	0.8232	39.767	S91	13.815	1.189	9.358	135.9204	32.0354	12.515	0.82	0.7392	206.392
S131	6.855	1.025	3.284	9.06	11.88	1.36	0.21	0.56	34.234	S92	14.15	1.168	10.418	181.0776	41.5076	13.62	0.83	1.5848	264.356
S132	6.23	1.012	2.55	62.364	20.734	1.455	0.08	1.272	95.697	S93	13.51	1.085	10.356	177.5292	39.9742	13.325	0.45	1.4016	257.631
S133	5.76	0.893	1.768	50.232	10.272	1.805	0.12	0.556	71.406	S94	13.565	1.31	10.418	212.8404	57.3104	13.565	0.61	2.9792	312.598
S134	5.46	0.891	3.996	42.3948	11.2398	3.295	0.16	0.3504	67.787	S95	13.585	1.29	11.144	114.4956	23.1206	13.635	0.1	0.1488	177.519
S135	7.08	0.74	3.126	16.422	13.457	3.24	0.12	0.476	44.661	S96	10.545	1.344	9.408	220.3608	25.9058	13.835	0.44	0.3584	282.197
S136	7.98	0.653	3.386	47.574	18.799	3.785	0.14	0.872	83.189	S97	13.925	1.355	9.974	221.8992	28.0792	13.475	0.48	0.3616	289.549
S137	3.895	0.657	3.234	21.7776	12.3876	3.755	0.28	0.4648	46.451	S98	13.915	1.257	9.962	165.9192	26.8642	13.805	0.45	0.2816	232.454
S138	6.38	0.782	2.458	59.5032	14.8982	3.11	0.26	0.8536	88.245	S99	14.8	1.272	9.188	154.6824	65.5774	13.57	0.57	3.5352	263.195
S139	6.665	0.681	1.7	12.522	15.177	2.465	0.12	0.956	40.286	S100	13.605	1.009	15.878	221.952	27.342	15.11	0.78	0.476	296.152
S140	6.045	0.621	5.464	43.7484	9.4884	2.225	0.19	0.5032	68.285	S101	13.62	1.115	16.57	126.0852	29.8402	14.765	0.55	0.6896	203.235
S141	4.89	0.813	3.92	13.0884	15.5634	1.765	0.13	0.8232	40.993	S102	13.355	1.088	18.134	67.1352	25.7552	14.265	0.69	0.3696	140.792
S142	6.455	0.83	2.51	50.91	11.49	1.325	0.21	0.56	74.29	S103	13.205	1.548	20.466	140.382	37.862	12.64	0.99	1.356	228.449
S143	7.39	0.887	2.412	32.514	23.284	1.295	0.08	1.272	69.134	S104	13.21	1.453	18.794	238.1304	43.5504	16.435	0.82	1.6992	334.092
S144	6.46	0.679	3.12	54.372	9.837	1.915	0.12	0.556	77.059	S105	19.46	1.476	18.706	194.8968	41.2568	13.955	0.86	1.4664	292.077
S145	4.735	0.511	1.504	10.0548	8.5648	1.025	0.16	0.3504	26.905	S106	19.265	1.609	15.804	140.5524	30.2274	13.715	0.98	0.6552	222.808
S146	6.925	0.533	2.858	38.412	13.687	2.565	0.12	0.476	65.576	S107	16.15	1.63	16.278	92.5776	45.2676	13.95	0.82	2.0648	188.738
S147	4.795	0.568	2.928	59.574	16.879	3.515	0.14	0.872	89.271	S108	16.155	1.211	17.218	76.3884	37.3034	14.62	0.92	1.3032	165.119
S148	4.685	0.447	1.742	9.1176	11.7526	3.1	0.28	0.4648	31.589	S109	17.365	1.304	18.756	62.2632	35.6632	21.075	0.99	0.9336	158.35
S149	3.415	0.399	2.434	49.4832	14.6782	1.73	0.26	0.8536	73.253	S110	16.43	1.303	10.978	65.1876	27.6976	14.52	0.88	0.6248	137.621
S150	3.67	0.613	1.928	29.232	9.607	1.765	0.12	0.556	47.491	S111	13.635	1.508	18.578	91.5084	32.0784	4.14	0.92	0.7432	163.111
										S112	13.945	1.025	19.642	59.0724	31.1324	5.38	0.67	0.6552	131.522
										S113	13.965	0.976	19.594	60.6804	32.0304	6.675	0.82	0.7392	135.48



Appendix table 5 Summary statistics of the Linear Structural Equation Model (LSEM) for the Cu, linking soil quality, water quality, agricultural activities, anthropogenic pressure and climatic effect. Significant effects (p<0.05).

		Direct effect				
Dependent	independent	estimate	SE	β	Z value	P values
Cu	Anthro pressure	3.95531	0.54803	0.48599	7.2173	0.001
Cu	Temp	-71.58638	103.99203	-0.0455	-0.6884	0.491
Cu	Precipitation	2.74706	4.18537	0.04258	0.6563	0.512
Cu	Soil quality	-0.03481	1.86746	-0.0012	-3.0186	0.025
Cu	water quality	3.04932	1.05695	0.7591	2.8659	0.007
Cu	agricultural activity	0.11033	0.02528	0.2877	4.3645	0.040
Soil quality	Temp	8.85556	4.38002	0.15411	2.0218	0.043
Soil quality	Precipitation	-0.16921	0.18235	-0.0717	-0.9280	0.353
Soil quality	water quality	0.00975	0.00236	0.31975	4.1364	0.001
Soil quality	agricultural activity	0.00112	0.00107	0.07991	1.0487	0.294
water quality	Temp	-38.83357	155.25927	-0.0206	-0.2501	0.802
water quality	Precipitation	12.98152	6.22717	0.16791	2.0847	0.037
water quality	Anthro pressure	0.55069	0.80272	0.05647	0.6860	0.493
agricultural activity	Temp	-154.737	332.51258	-0.0377	-0.4654	0.642
agricultural activity	Anthro pressure	5.35620	1.72150	0.25238	3.1114	0.002
agricultural activity	Precipitation	1.71091	13.52547	0.01017	0.1265	0.899
agricultural activity	water quality	0.06304	0.17483	0.02897	0.3606	0.718
Indirect effect						
Anthro, , Cu	b3*b5	0.027	0.051	0.003	4.538	0.001
Anthro, , S, Cu	b1*b4*b5	-0.000	0.010	-0.000	-2.019	0.085
Anthro, AA, S, Cu	b2*b4*b5	0.004	0.012	0.000	2.318	0.050
Temp, S, Cu	b1*b7	-0.000	0.000	-0.000	-2.019	0.085
Preci, S, Cu	b2*b7	0.591	0.233	0.073	2.534	0.011
Temp, , S, Cu	b2*b4*b5	-0.000	0.011	-0.000	-2.019	0.085
Preci, , S, Cu	b3*b4*b5	-0.308	16.538	-0.000	-2.019	0.085
Total effect	b1*b2*b3*b4*b5*b6*b7	-1.915	7.970	-0.001	-3.240	0.010
Variables	R <sup>2</sup>					
Soil quality	0.31					
ater quality	0.33					
Agricultural activity	0.32					
Cu	0.39					

Appendix table 6 Summary statistics of the Structural Equation Model fit values for Cu.

Model fit indices	values
Chi squ	76.00
P value	0.054
Comparative Fit Index (CFI)	0.838
Tucker-Lewis Index (TLI)	0.920
Relative Noncentrality Index (RNI)	0.019
Bollen's Incremental Fit Index (IFI)	0.888
GFI	0.944
RMSEA	0.107
SRMR	0.104
AIC	4160
BIC	4236



Appendix table 7 Summary statistics of the Linear Structural Equation Model (LSEM) for Cr, linking with soil quality, water quality, agricultural activities, anthropogenic pressure and climatic effects in Aik stream area. Significant effects ( $p < 0.05$ ).

Direct effect						
Dependent	independent	estimate	SE	$\beta$	Z value	P values
Cr	Anthro pressure	5.39092	1.71955	0.25402	3.135	0.002
Cr	Temp	23.07368	3.31432	0.60261	6.962	0.001
Cr	Precipitation	-874.016	497.89678	-0.1181	-1.755	0.079
Cr	Soil quality	-19.85901	20.41698	-0.0654	-0.973	0.331
Cr	water quality	4.35931	11.62132	0.8335	3.375	0.008
Cr	agricultural activity	0.72746	0.28062	0.18528	2.592	0.210
Soil quality	Temp	-0.22175	0.01858	0.62726	9.960	0.001
Soil quality	Precipitation	0.00889	3.48373	0.02221	0.364	0.716
Soil quality	water quality	-9.75e-4	0.14168	-0.0948	-1.565	0.118
Soil quality	agricultural activity	-17.16259	0.00183	0.29404	4.862	0.001
water quality	Temp	13.11099	8.55e-4	-0.0701	-1.140	0.254
water quality	Precipitation	0.01918	151.72849	-0.0091	-0.113	0.910
water quality	Anthro pressure	-157.185	6.22553	0.16959	2.106	0.035
agricultural activity	Temp	2.52928	0.03695	0.04173	0.519	0.604
agricultural activity	Anthro pressure	5.39092	332.58731	-0.0383	-0.473	0.636
agricultural activity	Precipitation	2.52928	13.33947	0.01503	0.190	0.850
agricultural activity	water quality	5.39092	1.71955	0.25402	3.135	0.002
Indirect effect						
Anthro, , C r	b3*b5	-0.807	2.152	0.001	0.365	0.723
Anthro, , S , Cr	b1*b4*b5	0.023	0.065	-0.000	-0.368	0.763
Anthro, AA, S , Cr	b2*b4*b5	-0.004	0.013	0.002	1.635	0.615
Temp, S , Cr	b1*b7	0.075	0.150	-0.001	0.167	0.794
Preci, S , Cr	b2*b7	-5.520	21.147	0.000	-0.161	0.914
Temp, , S , Cr	b2*b4*b5	0.665	6.146	-0.002	0.178	0.910
Preci, , S , Cr	b3*b4*b5	-12.485	110.481	-0.000	-0.374	0.776
Total effect	b1*b2*b3*b4*b5*b6*b7	-0.668	2.347	0.023	0.356	0.798
Variables	R <sup>2</sup>					
Soil quality	0.46					
ater quality	0.31					
Agricultural activity	0.24					
Cr	0.35					

Appendix table 8 Summary statistics of the Structural Equation Model fit values for Cr.

Model fit indices	values
Chi squ	9.01
P value	0.064
Comparative Fit Index (CFI)	0.989
Tucker-Lewis Index (TLI)	0.902
Relative Noncentrality Index (RNI)	0.049
Bollen's Incremental Fit Index (IFI)	0.982
GFI	0.999
RMSEA	0.153
SRMR	0.033
AIC	4557
BIC	4629



Appendix table 9 Summary statistics of the Linear Structural Equation Model (LSEM) for Zn, linking with soil quality, water quality, agricultural activities, anthropogenic pressure and climatic effects in Aik stream area. Significant effects ( $p < 0.05$ ).

Direct effect						
Dependent	independent	estimate	SE	$\beta$	Z value	P values
Zn	Anthro pressure	5.95995	0.80966	0.48599	7.361	0.001
Zn	Temp	14.23051	117.84790	-0.0455	0.121	0.904
Zn	Precipitation	2.03129	4.82955	0.04258	0.421	0.674
Zn	Soil quality	-2.32296	2.76083	-0.0012	-8.841	0.000
Zn	water quality	3.00724	1.06655	0.65910	2.109	0.003
Zn	agricultural activity	0.09253	0.02905	0.28770	3.186	0.001
Soil quality	Temp	0.18503	0.01858	0.62726	9.960	0.001
Soil quality	Precipitation	1.26635	3.48373	0.02221	0.364	0.716
Soil quality	water quality	-0.22175	0.14168	-0.0948	-1.565	0.118
Soil quality	agricultural activity	0.00889	0.00183	0.29404	4.862	0.001
water quality	Temp	-9.754	8.55e-4	-0.0701	-1.140	0.254
water quality	Precipitation	-17.19186	151.72850	-0.0091	-0.113	0.910
water quality	Anthro pressure	13.11089	6.22553	0.16959	2.106	0.035
agricultural activity	Temp	0.01918	0.03695	0.04173	0.519	0.604
agricultural activity	Anthro pressure	-157.1855	332.58731	-0.0383	-0.473	0.636
agricultural activity	Precipitation	2.52928	13.33947	0.01503	0.190	0.850
agricultural activity	water quality	5.39092	1.71955	0.25402	3.135	0.002
Indirect effect						
Anthro, , Z n	b3*b5	-0.430	0.513	-0.044	-5.838	0.002
Anthro, , S , Zn	b1*b4*b5	0.012	0.018	0.001	0.662	0.508
Anthro, AA, S , Zn	b2*b4*b5	-0.002	0.005	-0.000	-0.436	0.663
Temp, S , Zn	b1*b7	-0.001	0.007	-0.000	-0.106	0.915
Preci, S , Zn	b2*b7	0.499	0.223	0.052	2.234	0.025
Temp, , S , Zn	b2*b4*b5	-2.942	8.815	-0.002	-0.334	0.739
Preci, , S , Zn	b3*b4*b5	0.355	3.164	0.234	3.112	0.011
Total effect	b1*b2*b3*b4*b5*b6*b7	0.124	1.586	0.023	2.078	0.037
Variables	R <sup>2</sup>					
Soil quality	0.46					
ater quality	0.31					
Agricultural activity	0.24					
Zn	0.43					

Appendix table 10 Summary statistics of the Structural Equation Model fit values for Zn.

Model fit indices	values
Chi squ	0.333
P value	0.565
Comparative Fit Index (CFI)	0.954
Tucker-Lewis Index (TLI)	0.897
Relative Noncentrality Index (RNI)	0.039
Bollen's Incremental Fit Index (IFI)	0.932
GFI	0.912
RMSEA	0.001
SRMR	0.010
AIC	4126
BIC	4201



Appendix table 11 Summary statistics of the Linear Structural Equation Model (LSEM) for Cd, linking with soil quality, water quality, agricultural activities, anthropogenic pressure and climatic effects in Aik stream area. Significant effects ( $p < 0.05$ ).

		Direct effect				
Dependent	independent	estimate	SE	$\beta$	Z value	P values
Cd	Anthro pressure	0.46012	0.05105	0.66879	9.013	< .001
Cd	Temp	-14.40007	7.43061	-0.1083	-1.938	0.053
Cd	Precipitation	0.31783	0.30452	0.05828	1.044	0.297
Cd	Soil quality	0.25641	0.17408	0.10994	1.473	0.141
Cd	water quality	4.2374	1.02420	0.8730	3.102	0.002
Cd	agricultural activity	0.00180	0.00183	0.05554	0.983	0.326
Soil quality	Temp	0.18503	0.01858	0.62726	9.960	< .001
Soil quality	Precipitation	1.26635	3.48373	0.02221	0.364	0.716
Soil quality	water quality	-0.22175	0.14168	-0.0948	-1.565	0.118
Soil quality	agricultural activity	0.00889	0.00183	0.29404	4.862	< .001
water quality	Temp	-9.754	8.55e-4	-0.0701	-1.140	0.254
water quality	Precipitation	-17.16259	151.72850	-0.0091	-0.113	0.910
water quality	Anthro pressure	13.11099	6.22553	0.16959	2.106	0.035
agricultural activity	Temp	0.01918	0.03695	0.04173	0.519	0.604
agricultural activity	Anthro pressure	-157.1856	332.58731	-0.0383	-0.473	0.636
agricultural activity	Precipitation	2.52928	13.33947	0.01503	0.190	0.850
agricultural activity	water quality	5.39092	1.71955	0.25402	3.135	0.002
		Indirect effect				
Anthro, , N i	b3*b5	0.047	0.033	0.069	1.457	0.145
Anthro, , S , Ni	b1*b4*b5	-0.001	0.002	-0.002	-0.866	0.386
Anthro, AA, S , Ni	b2*b4*b5	0.000	0.000	0.000	0.481	0.630
Temp, S , Ni	b1*b7	-0.000	0.000	-0.000	-0.100	0.921
Preci, S , Ni	b2*b7	0.010	0.010	0.014	0.938	0.348
Temp, , S , Ni	b2*b4*b5	0.325	0.920	0.002	0.353	0.724
Preci, , S , Ni	b3*b4*b5	-0.039	0.347	-0.000	-0.113	0.910
Total effect	b1*b2*b3*b4*b5*b6*b7	0.007	0.097	0.000	0.076	0.940
Variables	R <sup>2</sup>					
Soil quality	0.46					
ater quality	0.34					
Agricultural activity	0.24					
Cd	0.53					

Appendix table 12 Summary statistics of the Structural Equation Model fit values for Cd.

Model fit indices	values
Chi squ	6.08
P value	0.299
Comparative Fit Index (CFI)	0.998
Tucker-Lewis Index (TLI)	0.992
Relative Noncentrality Index (RNI)	0.029
Bollen's Incremental Fit Index (IFI)	0.974
GFI	0.934
RMSEA	0.038
SRMR	0.027
AIC	3295
BIC	3365



Appendix table 13 Summary statistics of the Linear Structural Equation Model (LSEM) for Ni, linking with soil quality, water quality, agricultural activities, anthropogenic pressure and climatic effects in Aik stream area. Significant effects ( $p < 0.05$ ).

		Direct effect				
Dependent	independent	estimate	SE	$\beta$	Z value	P values
Ni	Anthro pressure	2.65845	0.38160	0.58072	3.135	< .001
Ni	Temp	-114.9320	55.54198	-0.1299	6.967	0.039
Ni	Precipitation	-0.65890	2.27618	-0.0181	-2.069	0.772
Ni	Soil quality	0.68356	1.30119	0.04405	-0.289	0.599
Ni	water quality	2.0295	1.03137	0.6631	1.525	0.001
Ni	agricultural activity	0.03935	0.01369	0.18240	-0.945	0.304
Soil quality	Temp	0.18503	0.01858	0.62726	2.874	< .001
Soil quality	Precipitation	1.26635	3.48373	0.02221	9.960	0.716
Soil quality	water quality	-0.22175	0.14168	-0.0948	0.364	0.118
Soil quality	agricultural activity	0.00889	0.00183	0.29404	-1.565	< .001
water quality	Temp	-9.754	8.55e-4	-0.0701	4.862	0.254
water quality	Precipitation	-17.16259	151.72850	-0.0091	-1.140	0.910
water quality	Anthro pressure	13.11099	6.22553	0.16959	-0.113	0.035
agricultural activity	Temp	0.01918	0.03695	0.04173	2.106	0.604
agricultural activity	Anthro pressure	-157.1855	332.58731	-0.0383	0.519	0.636
agricultural activity	Precipitation	2.52928	13.33947	0.01503	-0.473	0.850
agricultural activity	water quality	5.39092	1.71955	0.25402	3.190	0.002
		Indirect effect				
Anthro, , N i	b3*b5	0.126	0.241	0.028	0.525	0.600
Anthro, , S , Ni	b1*b4*b5	-0.004	0.008	-0.001	-0.472	0.637
Anthro, AA, S , Ni	b2*b4*b5	0.001	0.002	0.000	0.366	0.715
Temp, S , Ni	b1*b7	-0.003	0.007	-0.001	-0.450	0.653
Preci, S , Ni	b2*b7	0.212	0.100	0.046	2.119	0.034
Temp, , S , Ni	b2*b4*b5	0.866	2.896	0.001	0.299	0.765
Preci, , S , Ni	b3*b4*b5	-0.104	0.944	-0.000	-0.111	0.912
Total effect	b1*b2*b3*b4*b5*b6*b7	0.509	4.530	0.001	0.112	0.911
Variables	R <sup>2</sup>					
Soil quality	0.46					
ater quality	0.31					
Agricultural activity	0.24					
Ni	0.38					

Appendix Table 14 Summary statistics of the Structural Equation Model fit values for Ni.

Model fit indices	values
Chi squ	13.1
P value	0.023
Comparative Fit Index (CFI)	0.983
Tucker-Lewis Index (TLI)	0.938
Relative Noncentrality Index (RNI)	0.039
Bollen's Incremental Fit Index (IFI)	0.964
GFI	0.999
RMSEA	0.104
SRMR	0.038
AIC	3905
BIC	3968



Appendix table 15 Summary statistics of the Linear Structural Equation Model (LSEM) for Pb, linking with soil quality, water quality, agricultural activities, anthropogenic pressure and climatic effects in Aik stream area. Significant effects ( $p < 0.05$ ).

Direct effect						
Dependent	independent	estimate	SE	$\beta$	Z value	P values
Pb	Anthro pressure	7.62422	1.03242	0.56842	7.3848	< .001
Pb	Temp	-148.355	150.26963	-0.0572	-0.9873	0.324
Pb	Precipitation	11.31598	6.15824	0.10642	1.8375	0.066
Pb	Soil quality	0.26739	3.52038	0.00588	0.0760	0.939
Pb	water quality	2.12494	1.28486	0.89084	1.4722	0.001
Pb	agricultural activity	0.18758	0.03704	0.29681	5.0649	< .001
Soil quality	Temp	0.18503	0.01858	0.62726	9.9597	< .001
Soil quality	Precipitation	1.26635	3.48373	0.02221	0.3635	0.716
Soil quality	water quality	-0.22175	0.14168	-0.0948	-1.5652	0.118
Soil quality	agricultural activity	0.00889	0.00183	0.29404	4.8616	< .001
water quality	Temp	-9.754	8.554	-0.0701	-1.1396	0.254
water quality	Precipitation	-17.16259	151.72849	-0.0091	-0.1131	0.910
water quality	Anthro pressure	13.11099	6.22553	0.16959	2.1060	0.035
agricultural activity	Temp	0.01918	0.03695	0.04173	0.5190	0.604
agricultural activity	Anthro pressure	-157.1855	332.58731	-0.0383	-0.4726	0.636
agricultural activity	Precipitation	2.52928	13.33947	0.01503	0.1896	0.850
agricultural activity	water quality	5.39092	1.71955	0.25402	3.1351	0.002
Indirect effect						
Anthro, , P b	b3*b5	-0.807	2.152	0.001	0.365	0.723
Anthro, , S , Pb	b1*b4*b5	0.023	0.065	-0.000	-0.368	0.763
Anthro, AA, S , Pb	b2*b4*b5	-0.004	0.013	0.002	1.635	0.615
Temp, S , Pb	b1*b7	0.075	0.150	-0.001	0.167	0.794
Preci, S , Pb	b2*b7	-5.520	21.147	0.000	-0.161	0.914
Temp, , S , Pb	b2*b4*b5	0.665	6.146	-0.002	0.178	0.910
Preci, , S , Pb	b3*b4*b5	-12.485	110.481	-0.000	-0.374	0.776
Total effect	b1*b2*b3*b4*b5*b6*b7	-0.668	2.347	0.023	0.356	0.798
Variables	R <sup>2</sup>					
Soil quality	0.46					
ater quality	0.31					
Agricultural activity	0.24					
Pb	0.43					

Appendix table 16 Summary statistics of the Structural Equation Model fit values for Pb.

Model fit indices	values
Chi squ	24.00
P value	0.028
Comparative Fit Index (CFI)	0.957
Tucker-Lewis Index (TLI)	0.617
Relative Noncentrality Index (RNI)	0.029
Bollen's Incremental Fit Index (IFI)	0.994
GFI	0.991
RMSEA	0.270
SRMR	0.048
AIC	4221
BIC	4293



Appendix table 17 Summary statistics of the Linear Structural Equation Model (LSEM) for As, linking with soil quality, water quality, agricultural activities, anthropogenic pressure and climatic effects in Aik stream area. Significant effects ( $p < 0.05$ ).

		Direct effect				
Dependent	independent	estimate	SE	$\beta$	Z value	P values
As	Anthro pressure	0.17948	0.14417	0.05763	1.245	0.213
As	Temp	0.37409	0.08233	0.28466	4.544	< .001
As	Precipitation	0.00406	0.00198	0.10141	2.052	0.040
As	Soil quality	3.304	8.60e-4	0.01793	0.384	0.701
As	water quality	0.18598	0.01855	0.62521	10.028	< .001
As	agricultural activity	1.56888	3.47876	0.02724	0.451	0.652
Soil quality	Temp	-0.24452	0.14205	-0.1031	-1.721	0.085
Soil quality	Precipitation	0.00899	0.00183	0.29499	4.922	< .001
Soil quality	water quality	-9.854	8.52e-4	-0.0703	-1.155	0.248
Soil quality	agricultural activity	0.53729	0.80543	0.05504	0.667	0.505
water quality	Temp	-43.05111	155.96146	-0.0227	-0.276	0.783
water quality	Precipitation	13.28415	6.28052	0.17082	2.115	0.034
water quality	Anthro pressure	5.36221	1.72828	0.25239	3.103	0.002
agricultural activity	Temp	-152.7123	334.2446	-0.0371	-0.457	0.648
agricultural activity	Anthro pressure	1.55913	13.65701	0.00921	0.114	0.909
agricultural activity	Precipitation	0.06368	0.17553	0.02926	0.363	0.717
agricultural activity	water quality	0.06304	0.17483	0.02897	0.3606	0.718
		Indirect effect				
Anthro, , As	b3*b5	0.070	0.017	0.003	0.178	< .001
Anthro, , S , As	b1*b4*b5	0.002	0.003	-0.000	0.006	0.526
Anthro, AA, S , As	b2*b4*b5	0.002	0.003	0.000	0.005	0.513
Temp, S , As	b1*b7	0.000	0.000	-0.000	0.000	0.806
Preci, S , As	b2*b7	-0.000	0.000	0.073	-0.000	0.759
Temp, , S , As	b2*b4*b5	0.002	0.005	-0.000	0.005	0.703
Preci, , S , As	b3*b4*b5	-0.002	0.002	-0.000	-0.005	0.292
Total effect	b1*b2*b3*b4*b5*b6*b7	0.587	1.308	-0.001	0.008	0.654
Variables	R <sup>2</sup>					
Soil quality	0.46					
ater quality	0.34					
Agricultural activity	0.33					
Hg	0.61					

Appendix table 18 Summary statistics of the Structural Equation Model fit values for As.

Model fit indices	values
Chi squ	18.6
P value	0.078
Comparative Fit Index (CFI)	0.974
Tucker-Lewis Index (TLI)	0.883
Relative Noncentrality Index (RNI)	0.019
Bollen's Incremental Fit Index (IFI)	0.943
GFI	0.998
RMSEA	0.156
SRMR	0.014
AIC	3063
BIC	3129



Appendix table 19 Summary statistics of the Linear Structural Equation Model (LSEM) for Hg, linking with soil quality, water quality, agricultural activities, anthropogenic pressure and climatic effects in Aik stream area. Significant effects ( $p < 0.05$ ).

		Direct effect				
Dependent	independent	estimate	SE	$\beta$	Z value	P values
Hg	Anthro pressure	2.40501	0.01048	0.37934	4.0062	< .001
Hg	Temp	0.04201	1.52065	0.00470	0.0663	0.947
Hg	Precipitation	0.10084	0.06266	-0.0558	-0.7855	0.432
Hg	Soil quality	-0.04922	0.03579	0.10955	1.1395	0.254
Hg	water quality	0.04078	8.60e-4	0.00478	0.0630	0.950
Hg	agricultural activity	5.42e-5	3.74e-4	0.19355	2.6982	0.007
Soil quality	Temp	0.00101	0.14205	0.62521	10.0279	< .001
Soil quality	Precipitation	-0.24452	0.00183	0.02724	0.4510	0.652
Soil quality	water quality	0.00899	8.52e-4	-0.1031	-1.7214	0.085
Soil quality	agricultural activity	-9.85e-4	0.83105	0.29499	4.9221	< .001
water quality	Temp	0.46250	156.00571	-0.0703	-1.1555	0.248
water quality	Precipitation	-40.89675	6.27842	0.04738	0.5565	0.578
water quality	Anthro pressure	13.25082	0.03820	-0.0216	-0.2621	0.793
agricultural activity	Temp	0.01386	1.72646	0.17039	2.1105	0.035
agricultural activity	Anthro pressure	5.39642	334.30680	0.03016	0.3628	0.717
agricultural activity	Precipitation	-155.4536	13.46243	0.25400	3.1257	0.002
agricultural activity	water quality	2.40501	0.17483	-0.0377	-0.4650	0.642
		Indirect effect				
Anthro, , H g	b3*b5	0.027	0.051	0.003	4.538	0.001
Anthro, , S ,Hg	b1*b4*b5	-0.000	0.010	-0.000	-2.019	0.085
Anthro, AA, S , Hg	b2*b4*b5	0.004	0.012	0.000	2.318	0.050
Temp, S , Hg	b1*b7	-0.000	0.000	-0.000	-2.019	0.085
Preci, S , Hg	b2*b7	0.591	0.233	0.073	2.534	0.011
Temp, , S , Hg	b2*b4*b5	-0.000	0.011	-0.000	-2.019	0.085
Preci, , S , Hg	b3*b4*b5	-0.308	16.538	-0.000	-2.019	0.085
Total effect	b1*b2*b3*b4*b5*b6*b7	-1.915	7.970	-0.001	-3.240	0.010
Variables	R <sup>2</sup>					
Soil quality	0.48					
ater quality	0.34					
Agricultural activity	0.25					
Hg	0.24					

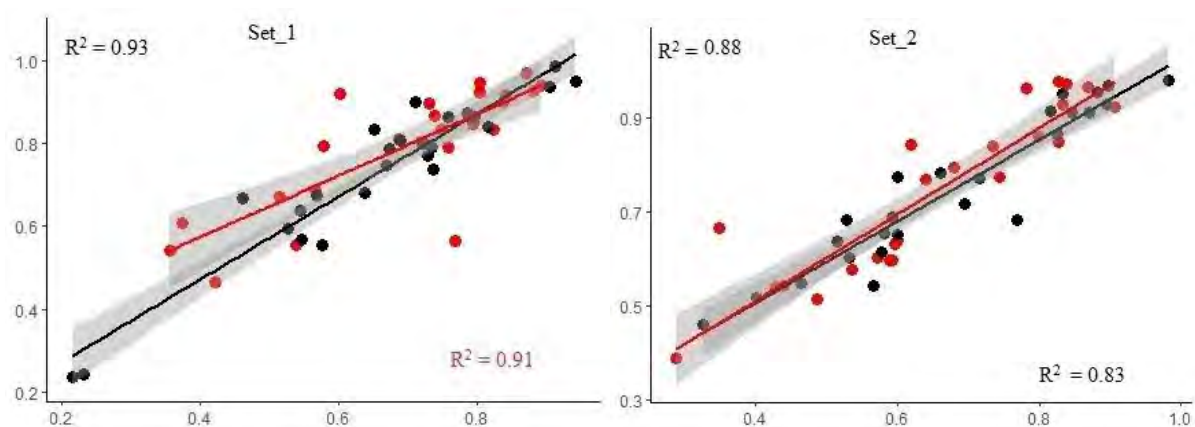
Appendix table 20 Summary statistics of the Structural Equation Model fit values for Hg.

Model fit indices	values
Chi squ	7.58
P value	0.108
Comparative Fit Index (CFI)	0.992
Tucker-Lewis Index (TLI)	0.963
Relative Noncentrality Index (RNI)	0.029
Bollen's Incremental Fit Index (IFI)	0.963
GFI	0.999
RMSEA	0.078
SRMR	0.032
AIC	2804
BIC	2870

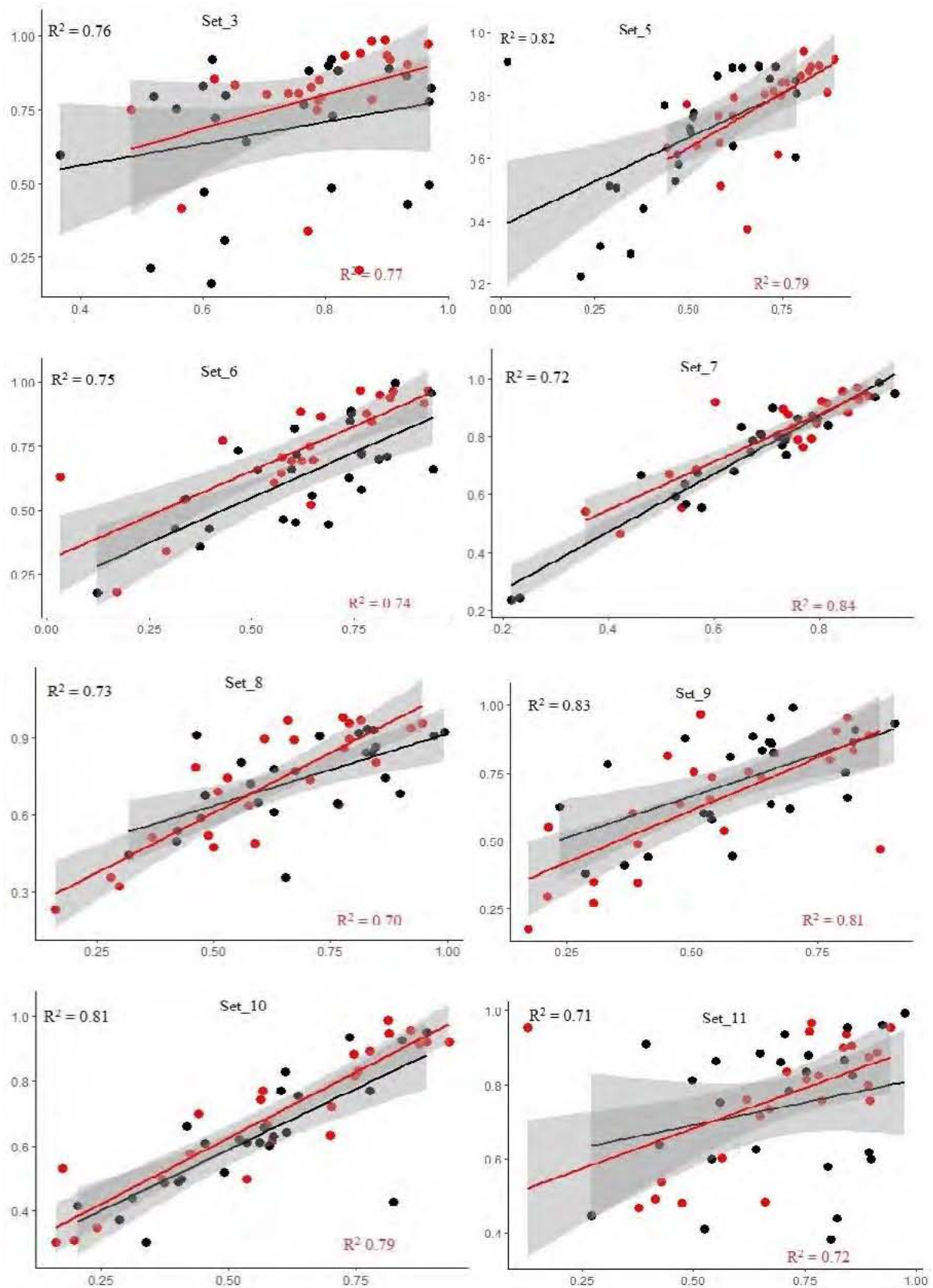


Appendix Table 21 Average NDVI and NDWI of the study area during (1998-2023).

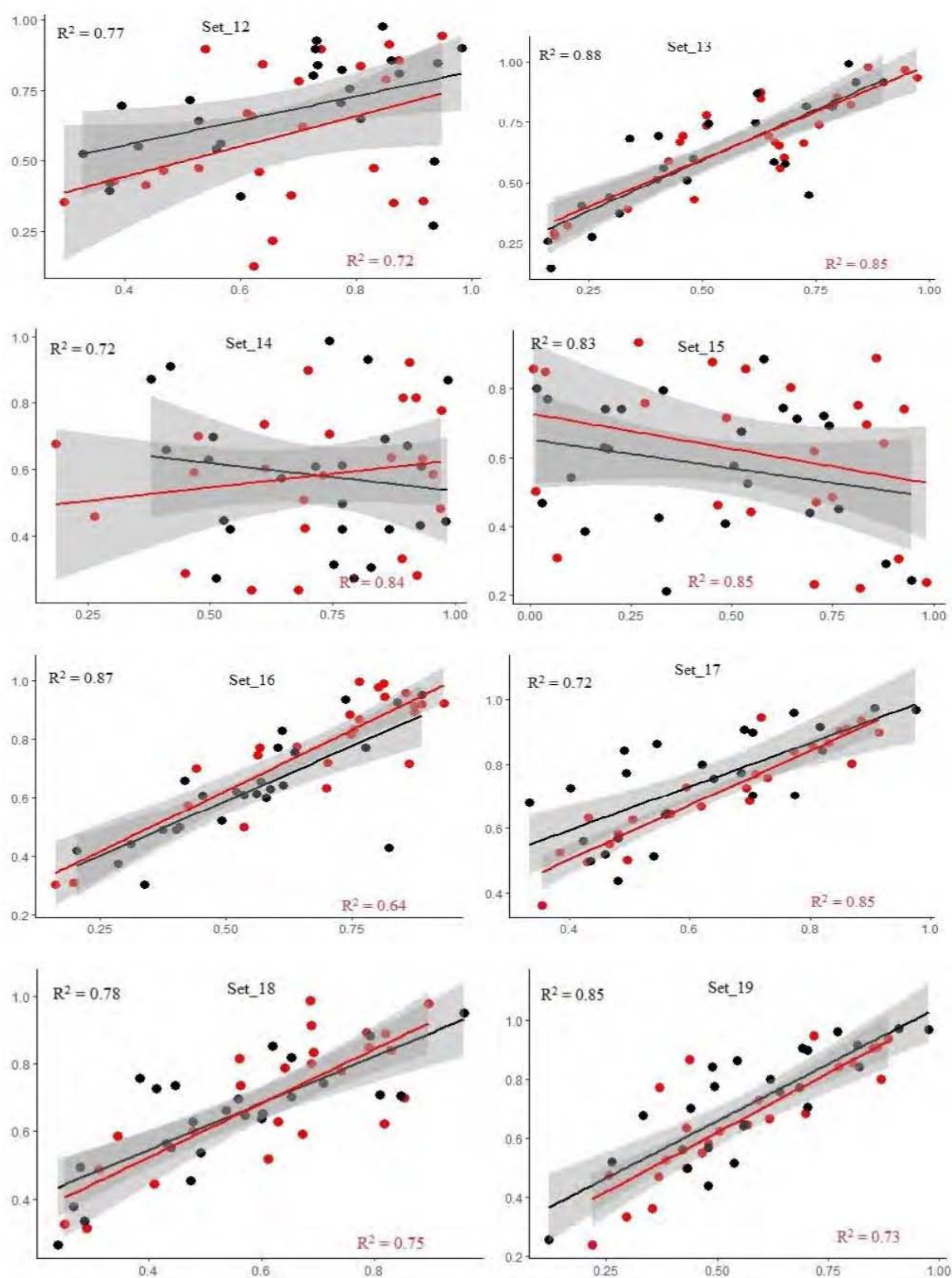
year	NDVI	NDWI
1998	0.785785	0.432887
1999	0.662864	0.534491
2000	0.785785	0.432887
2001	0.628638	0.449074
2002	0.638108	0.49074
2003	0.602823	0.3131
2004	0.793867	0.564105
2005	0.282339	0.53131
2006	0.386715	0.1048
2007	0.611049	0.313108
2008	0.3831	0.41048
2009	0.285867	0.471594
2010	0.287401	0.359115
2011	0.285666	0.471594
2012	0.240149	0.359115
2013	0.288717	0.471596
2014	0.187407	0.201062
2015	0.104789	0.210933
2016	0.102905	0.16361
2017	0.104789	0.133181
2018	0.102905	0.20933
2019	0.285868	0.21158
2020	0.28723	0.206609
2021	0.285868	0.1158
2022	0.28723	0.26609
2023	0.118723	0.2341





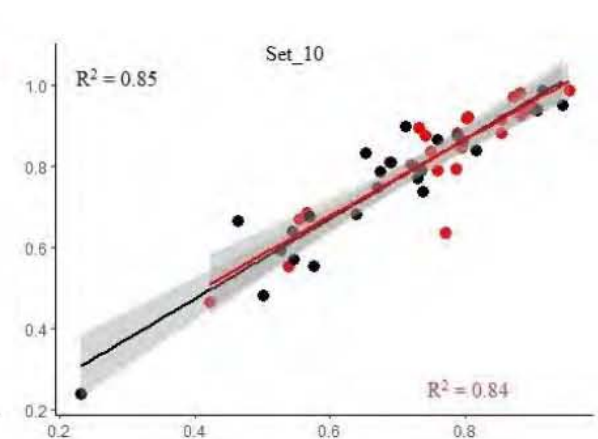
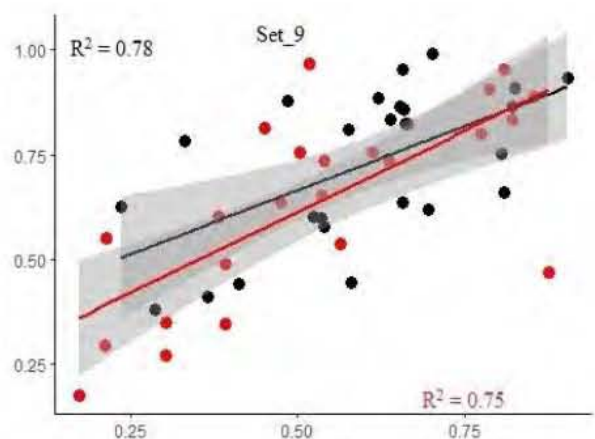
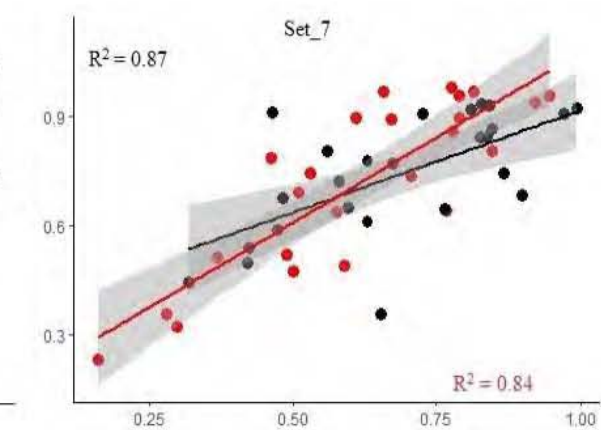
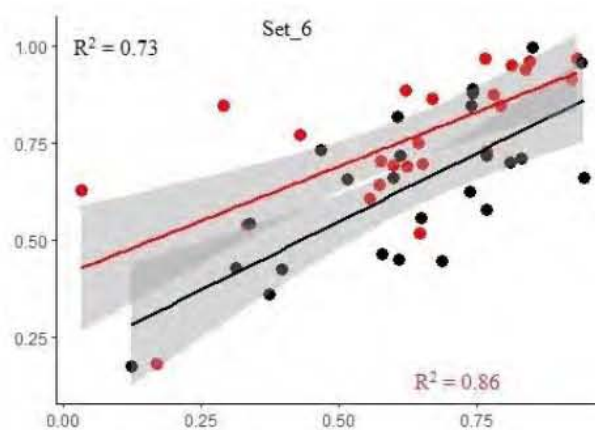
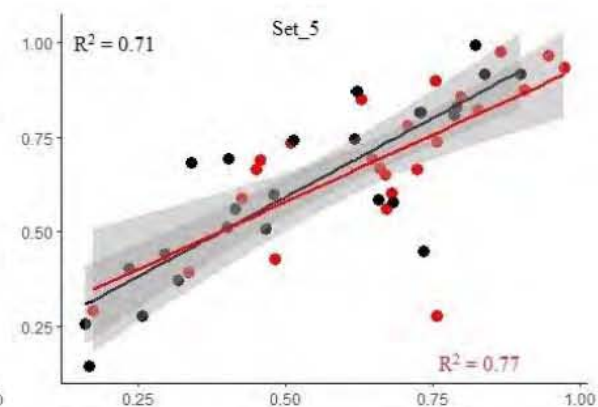
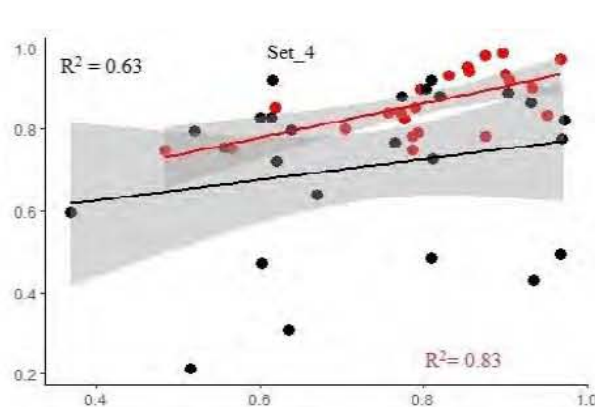
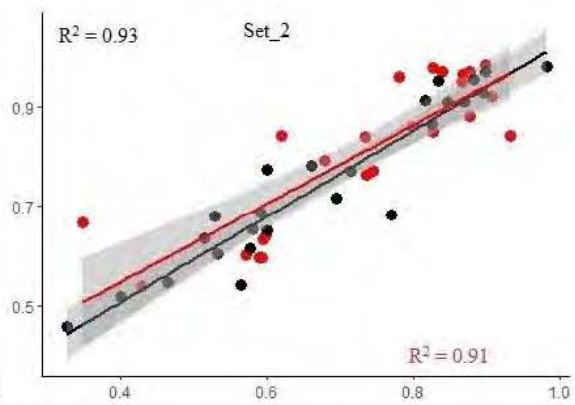
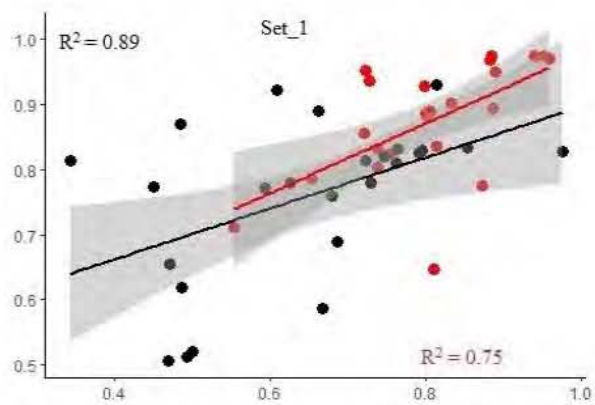




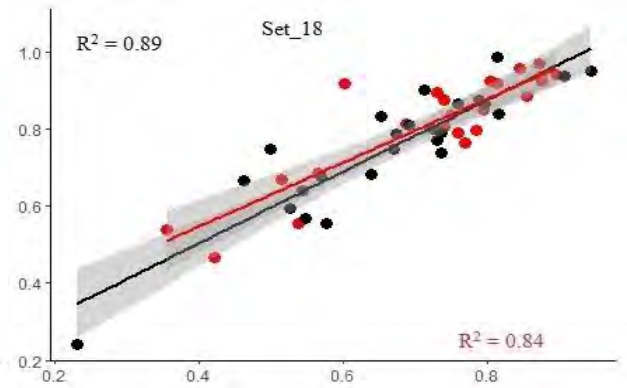
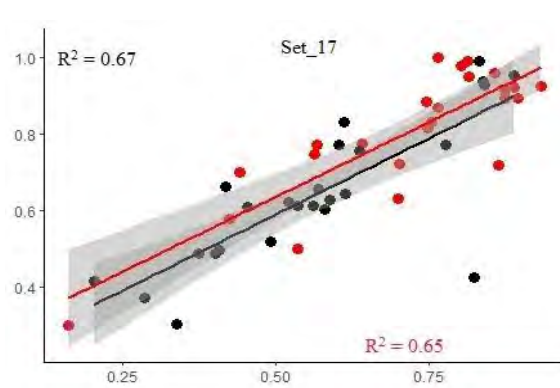
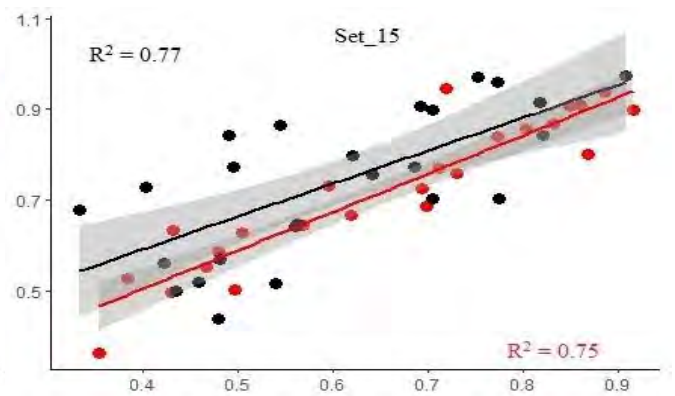
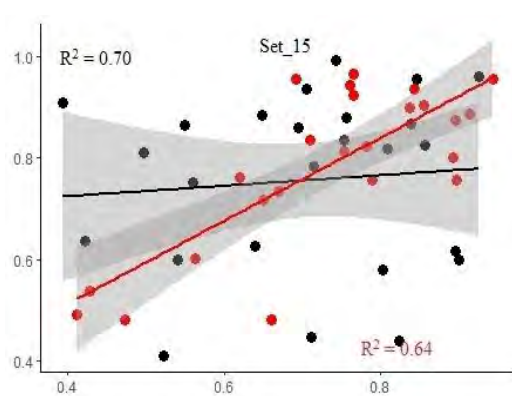
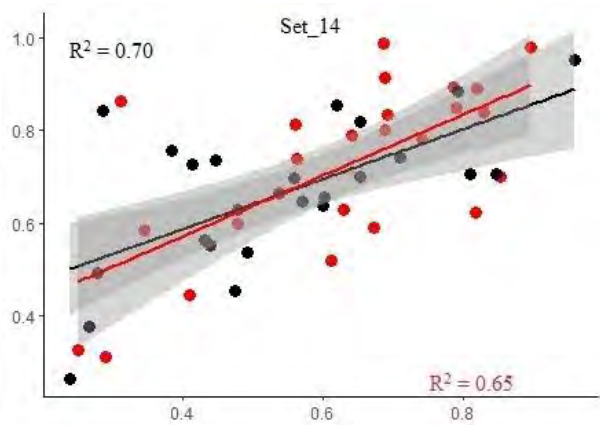
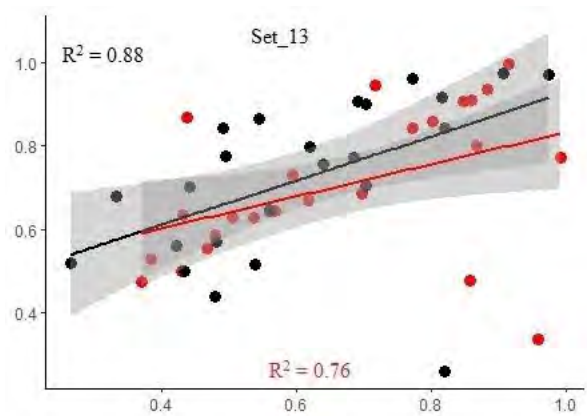
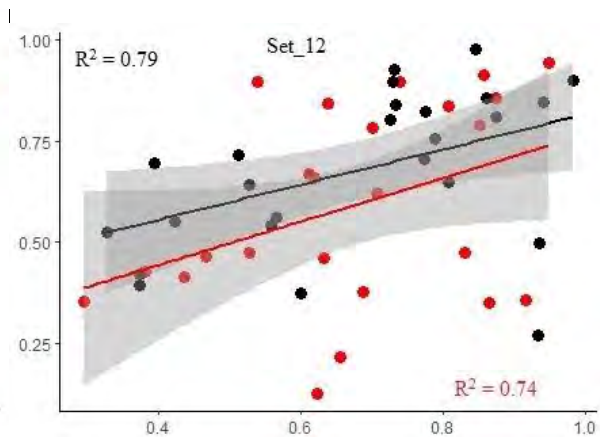
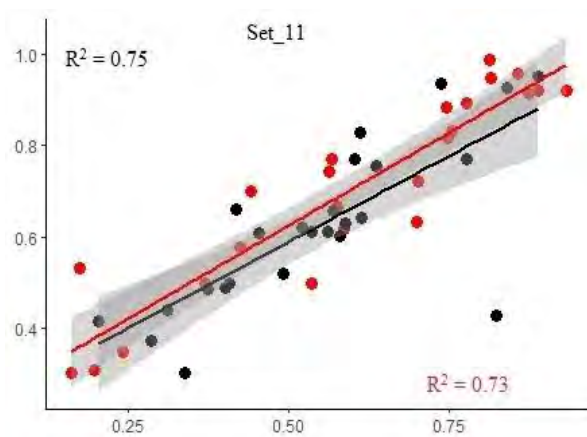


Appendix Figure 1 GB models for training and testing of the WQI in 19<sup>th</sup> different input combination, black line showed testing while red line indicate the training component.

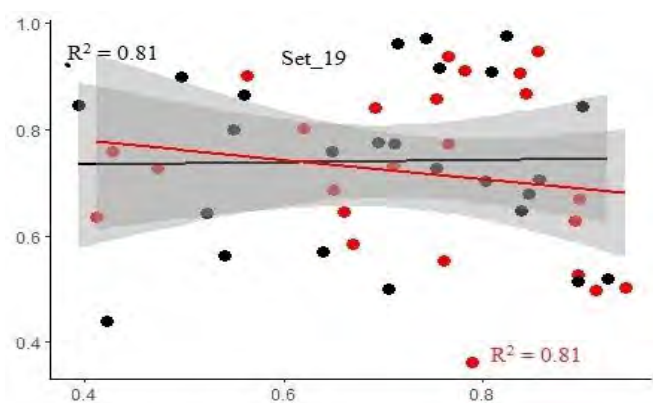












Appendix Figure 2 The graphical depiction of the Random Forest (RF) 19<sup>th</sup> input variables models for WQI Prediction, with training data shown in black and testing data in red.



Turnitin Originality Report

Ecological assessment of vegetation dynamics of the Aik stream with reference to the Industrial Pollution in Sialkot, Pakistan by Ujala Ejaz

From Master PhD (Master PhD)

- Processed on 04-Jul-2024 10:16 PKT
- ID: 2412371326
- Word Count: 80139

Similarity Index

12%

Similarity by Source

Internet Sources:

9%

Publications:

9%

Student Papers:

3%

**sources:**

- 1 < 1% match (student papers from 22-Feb-2010)  
[Submitted to Higher Education Commission Pakistan on 2010-02-22](#)
- 2 < 1% match (student papers from 22-Feb-2010)  
[Submitted to Higher Education Commission Pakistan on 2010-02-22](#)
- 3 < 1% match (student papers from 22-Feb-2010)  
[Submitted to Higher Education Commission Pakistan on 2010-02-22](#)
- 4 < 1% match (student papers from 22-Feb-2010)  
[Submitted to Higher Education Commission Pakistan on 2010-02-22](#)
- 5 < 1% match (student papers from 18-Jul-2019)  
[Submitted to Higher Education Commission Pakistan on 2019-07-18](#)
- 6 < 1% match (student papers from 24-Sep-2013)  
[Submitted to Higher Education Commission Pakistan on 2013-09-24](#)
- 7 < 1% match (student papers from 07-Mar-2018)  
[Submitted to Higher Education Commission Pakistan on 2018-03-07](#)
- 8 < 1% match (student papers from 16-Jun-2012)  
[Submitted to Higher Education Commission Pakistan on 2012-06-16](#)
- 9 < 1% match (student papers from 30-Jan-2013)  
[Submitted to Higher Education Commission Pakistan on 2013-01-30](#)
- 10 < 1% match (student papers from 06-Mar-2012)  
[Submitted to Higher Education Commission Pakistan on 2012-03-06](#)
- 11 < 1% match (student papers from 12-Jan-2023)  
[Submitted to Higher Education Commission Pakistan on 2023-01-12](#)
- 12 < 1% match (student papers from 22-Apr-2019)  
[Submitted to Higher Education Commission Pakistan on 2019-04-22](#)
- 13 < 1% match (student papers from 05-Jul-2019)  
[Submitted to Higher Education Commission Pakistan on 2019-07-05](#)
- 14 < 1% match (student papers from 27-Feb-2018)  
[Submitted to Higher Education Commission Pakistan on 2018-02-27](#)
- 15 < 1% match (student papers from 23-Jul-2013)  
[Submitted to Higher Education Commission Pakistan on 2013-07-23](#)

< 1% match (student papers from 31-Oct-2023)



