WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Assessing the Water Quality of Gilgit-Baltistan Province of Pakistan with Descriptive Statistics and Hierarchy Cluster Analysis

Malik Sajjad Mehmood [1]* and Shazmeena Komal

[1]*University of Engineering and Technology, Taxila, 47050, Pakistan*

[2]*Department of Thermofluid Science and Engineering, X'ian Jiaotong University China*

E-mail: msajjad.82@gmail.com

ABSTRACT

In this study, data on physical and chemical parameters related to the water quality of 94 samples from various parts of six districts of the Gilgit-Baltistan province of Pakistan were analyzed with the help of statistical analysis. Multivariate statistical analysis, including cluster analysis, principle component analysis, and descriptive analysis, was implemented in order to examine the similarities and differences in determining the spatial variations in the Indus River's water quality and major contaminant sources. Hierarchy cluster analysis (HCA) and descriptive analysis were implemented to examine the similarities and differences in the quality of water in all districts of the province. The data were used to analyze the water quality in the area and to gather important data on the number of monitoring stations that are necessary, the sources of pollution, and the assessment of geographical changes.

INTRODUCTION

Basin geography and geographic atmosphere along with climate change are generally influencing the quality of the river water and these changes are due to either by the nature and/or by anthropic consequences [1]. In addition to this, wastewater from industry, irrigation, and natural depletion, such as weathering, influence water quality thus making harmit ful for drinking purposes, for agriculture use, and other use [2]. Researchers used a variety of tools and methods to identify the potential causes of contaminations for different rivers around the world [3]. In this regard number of researchers have used analyzed the experimental data statistically and figured out the hidden patterns and responsible sources of contaminations to water quality [2, 4-6]. These methods include the study of the Hierarchy clusters, main feature analysis, and discriminatory analysis. These mathematical techniques enable estimating the multifaceted matrices of surface water content to be properly understood. Furthermore, these also allowed for figuring out the possible variables with an important impact on the water quality of the reservoir and are a viable tool for investigating the water flow and solving pollution problems.

This particular study aims at investigating the assessing the water quality of various districts of the Gilgit-Baltistan province of Pakistan and to highlight the contamination ions constituents. The hidden patterns/sources of pollution has also been identified with the help of hierarchical cluster analysis and discernment analysis

MATERIALS AND METHODS

STUDY AREA

Gilgit Baltistan is located at 35.8025° N, 74.983° E, and is also known as Pakistan's Northern Areas. Geographically, it covers 72,971 square kilometers. Afghanistan is to the north, China is to the east, and Kashmir is to the south of Gilgit Baltistan [7]. In 1970, the Northern Areas were renamed Gilgit Baltistan (GB), and later the Govt; of Pakistan has issued the GB Self Governance Order 2009 on August 29, 2009 which enables the residents of the Northern Areas (renamed Gilgit Baltistan) self-rule through a legally elected legislature. The population of the Gilgit Baltistan is estimated to be 1.5 to 2 million people [8]. The topographic map of GB provice is shown in figure 1 below.

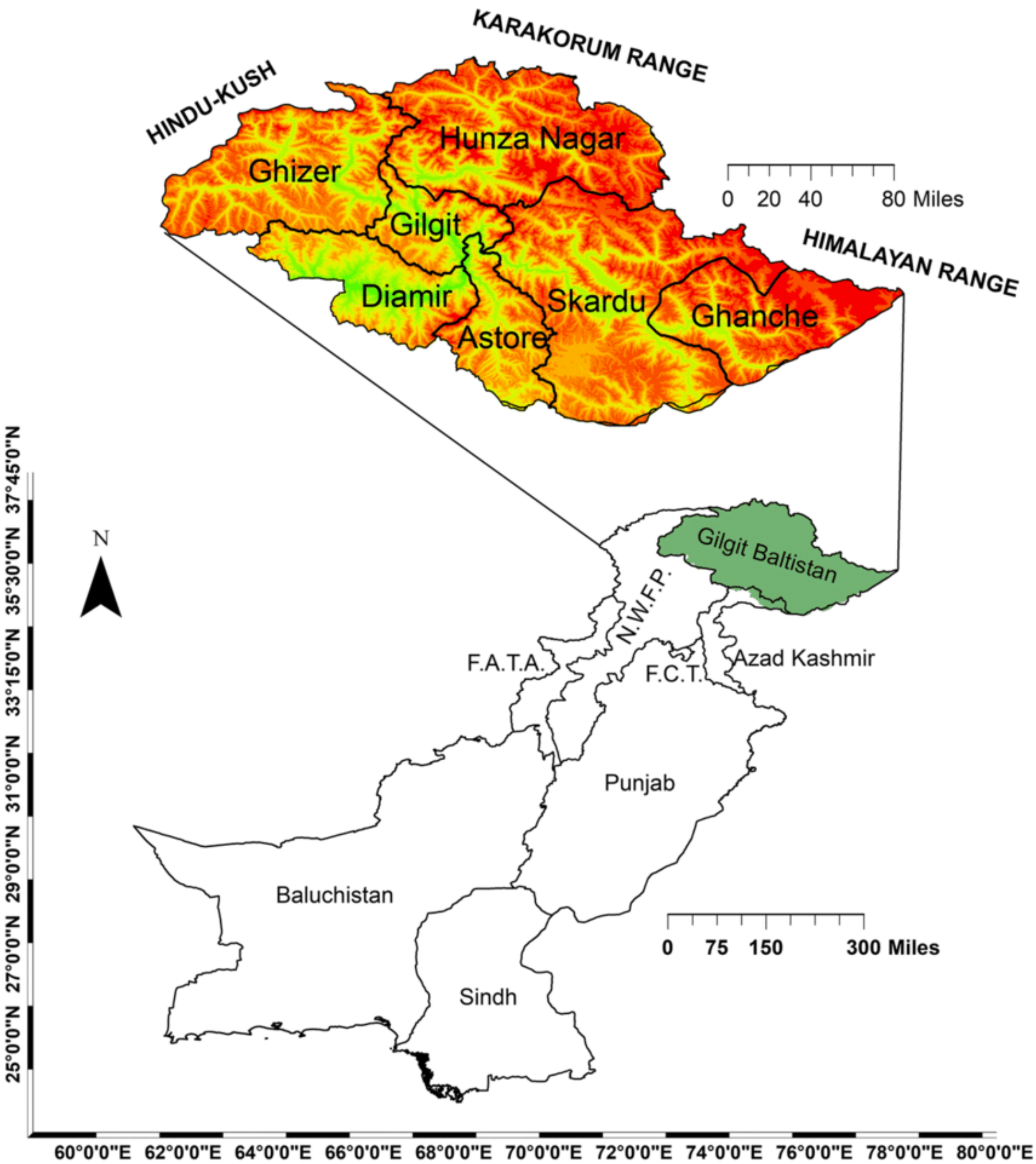

Figure 1: Topographic location of GB-Province on the map of Pakistan [9]

DATA COLLECTION FOR WATER QUALITY ASSESSMENT

The data of sixteen physicochemical parameters water quality (WQ) parameter were taken from the report [10] published by PCRWR which is the abbreviation of Govt; of Pakistan public office for conducting, managing, and regulations related to water research on water resources Pakistan Council of Research in Water Resources. These sixteen parameters include electrical Conductivity, pH, Turbidity, Calcium (Ca), Magnesium (Mg), Hard, Sodium (Na), Potassium (K), Chloride (Cl), Sulphate (SO4), Nitrate (NO3), Phosphorus(PO4), Total Dissolved Solids (TDS), Iron (Fe), Fluoride (F), Alkalinity (Alk) and Bicarbonate (HCO3). Afterward, multivariate statistical analysis is applied for classification and calculation of disorder values within the given data. The whole data has been divided in 6 regions which are labeled as G1, G2, G3, G4, G5 and G6, respectively. The details about all regions are given in table 1 below which are collected from the reference [10].

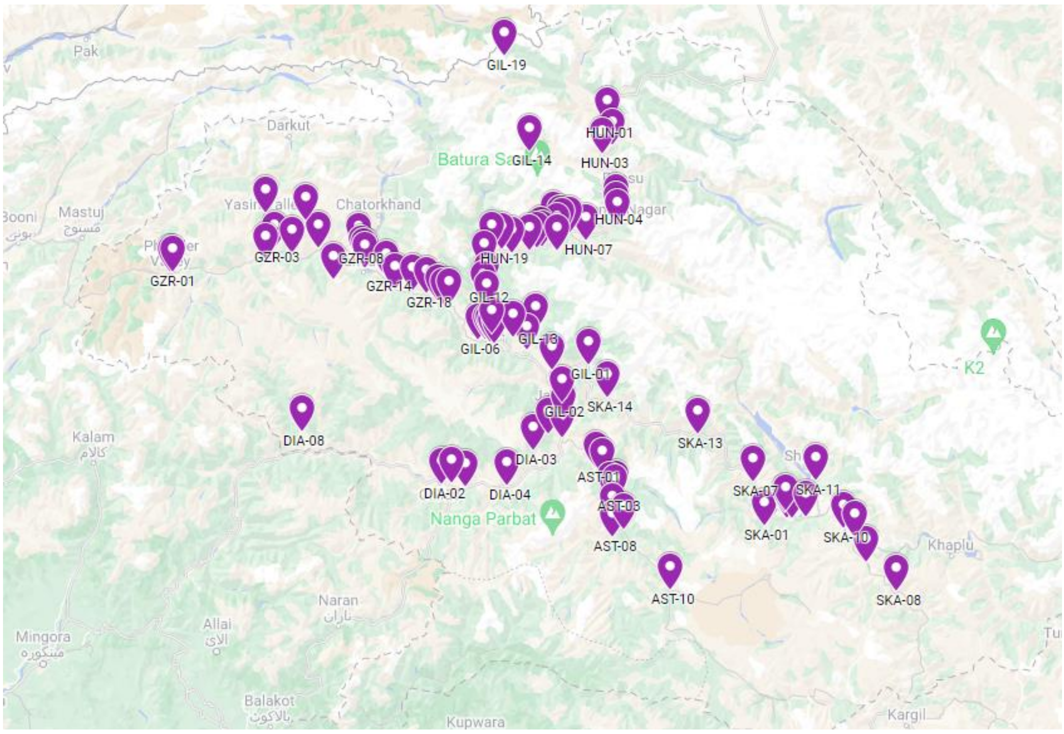

Figure 2: Location of sample points with sample codes

STATISTICAL ANALYSIS:

All the analysis were completed with the help of ORIGIN software (trial version) and Microsoft Excel 2007. The analysis includes the figuring out the coefficient of variance (CV), correlation analysis and Hierarchy cluster analysis (HCA). The details of HCA is given below

HIERARCHY CLUSTER ANALYSIS (HCA)

Cluster analysis is a multivariate database cluster whose primary function is to give tools for constructing objects based on unique parametric features. In cluster analysis, the basic approach for categorizing items is that one object is comparable to the other in terms of established descriptive criteria. To distinguish between important and irrelevant variables, the cluster object's results correlate to strong internal homogeneity and extremely high exterior heterogeneity. As a result, conceptual considerations must be used to justify variables in cluster analysis. Hierarchical agglomerative clustering is a common method for evaluating similarities between samples over a large data collection, and it is typically represented by a dendrogram. The major goal of this study is to determine the similarity and dissimilarity between each sample of monitoring sites by calculating a connection between individual clusters. The parametric mean value for each sample site is represented by a dendrogram. All the measured values are separated into three statistical categories.

1. Group A: this cluster could be considered less contaminated in general. Industrialization and urbanization should be kept to a bare minimum. As a result, the human impact should likewise be low.
2. Group B: this sort of cluster is typically found in very contaminated regions, such as industrial sites. Pollutants in industrial site samples often come from industrial wastewater treatment facilities, as well as sewage drainage from agricultural activities.
3. Group C: this group of sites might be classified as moderately polluted (MC). Nonpoint sources of pollution might contaminate these areas.

RESULTS AND DISCUSSION:

CORRELATION MATRIX OF WATER QUALITY PARAMETERS

The results of the correlation analysis are shown in Table 1 as a correlation matrix, which is created by adding the mean values of the quality parameters from the six Gilgit Baltistan areas (G1–G6). From the data, it can be shown that many measured parameters in the Gilgit–Baltistan area have a positive correlation.

Table 2: Correlation matrix of water quality parameters

	E.C	pH	Turbidity	HCO3 /AIK	Ca	Mg	Hard	Cl	Na	K	SO4	NO3	PO4	TDS	Fe	F
E.C	1.00	-0.07	-0.07	0.83	0.91	0.64	0.99	0.60	0.71	0.76	0.75	0.22	-0.09	1.00	0.05	0.54
pH	-0.07	1.00	0.12	-0.11	-0.16	0.07	-0.05	-0.24	-0.16	-0.10	0.09	-0.16	-0.11	-0.06	-0.01	-0.03
Turbidity	-0.07	0.12	1.00	-0.12	-0.07	-0.02	-0.06	-0.05	-0.10	0.04	-0.01	0.39	0.31	-0.07	0.41	-0.14
HCO3/AIK	0.83	-0.11	-0.12	1.00	0.80	0.42	0.80	0.63	0.72	0.58	0.37	0.31	-0.13	0.83	0.02	0.39
Ca	0.91	-0.16	-0.07	0.80	1.00	0.38	0.91	0.61	0.66	0.73	0.67	0.23	-0.05	0.91	0.12	0.53
Mg	0.64	0.07	-0.02	0.42	0.38	1.00	0.67	0.29	0.27	0.38	0.55	0.08	-0.14	0.64	-0.04	0.21
Hard	0.99	-0.05	-0.06	0.80	0.91	0.67	1.00	0.58	0.62	0.73	0.76	0.20	-0.10	0.99	0.06	0.49
Cl	0.60	-0.24	-0.05	0.63	0.61	0.29	0.58	1.00	0.61	0.51	0.33	0.46	-0.10	0.60	0.00	0.20
Na	0.71	-0.16	-0.10	0.72	0.66	0.27	0.62	0.61	1.00	0.56	0.36	0.28	-0.03	0.71	-0.03	0.61
K	0.76	-0.10	0.04	0.58	0.73	0.38	0.73	0.51	0.56	1.00	0.67	0.22	0.04	0.76	0.13	0.45
SO4	0.75	0.09	-0.01	0.37	0.67	0.55	0.76	0.33	0.36	0.67	1.00	-0.02	-0.06	0.75	0.03	0.41
NO3	0.22	-0.16	0.39	0.31	0.23	0.08	0.20	0.46	0.28	0.22	-0.02	1.00	0.29	0.21	0.19	-0.16
PO4	-0.09	-0.11	0.31	-0.13	-0.05	-0.14	-0.10	-0.10	-0.03	0.04	-0.06	0.29	1.00	-0.09	0.03	-0.08
TDS	1.00	-0.06	-0.07	0.83	0.91	0.64	0.99	0.60	0.71	0.76	0.75	0.21	-0.09	1.00	0.05	0.54
Fe	0.05	-0.01	0.41	0.02	0.12	-0.04	0.06	0.00	-0.03	0.13	0.03	0.19	0.03	0.05	1.00	0.11
F	0.54	-0.03	-0.14	0.39	0.53	0.21	0.49	0.20	0.61	0.45	0.41	-0.16	-0.08	0.54	0.11	1.00

COEFFICIENT OF VARIATION (CV)

Quality Parameters	G1	G2	G3	G4	G5	G6
E.C	0.804198	0.947753	90.01582	34.49165	23.65559	36.03856
pH	0.045908	0.022465	0.001276	0.00202	0.000476	0.00103
Turbidity	1.510006	1.186927	2.892771	0.535857	10.16127	7.950871
HCO3/AIK	0.608096	0.71492	12.52282	21.58193	1.737008	3.145204
Ca	0.734963	0.797677	0.384589	0.244489	0.153577	0.330335
Mg	0.987718	1.731726	0.144226	0.195322	0.332326	0.28376
Hard	0.8063	1.060042	16.04736	7.668284	4.398841	6.919221
Cl	1.258661	0	0.065754	0.058071	0.014644	0.035785
Na	1.554047	1.199794	1.798385	1.468688	3.309368	0.901364
K	0.928697	0.923197	0.925845	0.596717	0.899799	0.935604
SO4	1.318668	0.587619	8.997532	9.272623	7.660734	14.42815
NO3	0.508908	0.869009	0.078839	0.038086	0.033817	0.020951
PO4	0.130244	0.504178	0.197726	0.242216	1.137339	0.170052
TDS	0.818359	0.948198	1091.691	150.0261	134.5457	1081.157
Fe	1.522334	1.07638	0.001633	0	0.001215	0.001971
F	2.224763	1.156746	2.069845	0	3.717281	0.943502

Several sample sites in Gilgit Baltistan are organized into clusters based on similarities in water quality characteristics. However, by using this analytical method, it is possible to prepare for future events with the optimal number of sample stations. With greater knowledge of the dominant elements involved in the system under inquiry, this technique also helps in lowering the cost of monitoring.Figure 3 displays the dendrogram produced by a hierarchical cluster analysis that divided the data into three categories based on similar distinct features. Groups 1, 2, and 3 represent, respectively, the generally poor, fair, and excellent quality regions of Gilgit Baltistan.

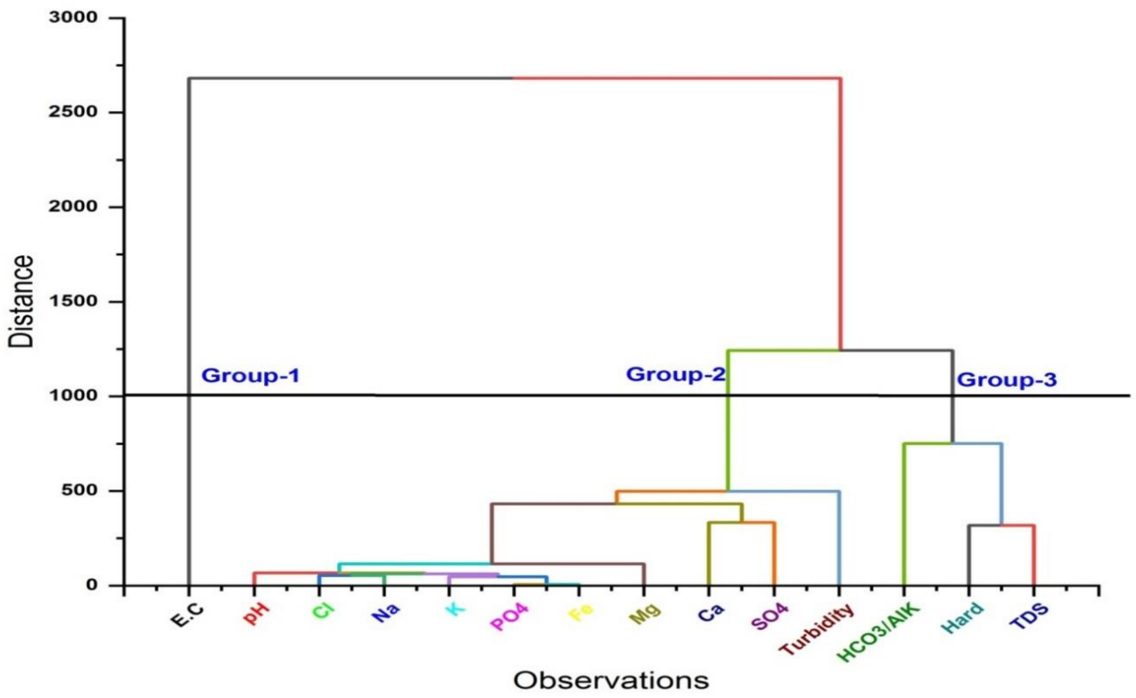

Figure 2: Dendrogram representing the cluster of monitoring stations

E.C, Ph, Cl, Na, K, PO4, and Fe formed group 1 and 2 in the aforementioned dendrogram, which is shown in Figure 2. The major factor for of contamination for sites blonging to group 1 areis E.C which can be linked to the water from group 1's snow-covered peaks and glaciers flows into the main sub-river through side nallas, and the majority of these nallas are clean and transparent with little to no domestic effluent.Group 2 is made up of Mg Ca SO4 Turbidity; analysis reveals that this group contains moderately contaminated water, with somewhat lower water quality than group 1. The population in these areas is mostly responsible for the decline in water quality, combined with the disposal of residential and commercial waste and human excreta effluent discharge sites. The respoonsible parameters for Group 3 are HCO3/AIK, Hard, and TDS; analysis reveals that this group is the most polluted group than groups 1 and 2 the reason behind that is given below;

- Greater water turbidity as a result of the presence of silt and rocks
- Illegally gathering medicinal plants and bushes, as well as indiscriminately felling trees and other important trees for the environment.
- A greater salinity percentage
- Discharge of human being waste effluent

CONCLUSION:

The water quality assessment of six districts in Gilgit Baltistan at the village level formed the basis for this study. Test results from all water sources were compared to the National Drinking Water Quality Standards' permissible limits for physicochemical, heavy metal, and microbiological parameters. The study revealed that just 22% of the sources provided safe water, while the remaining 78% were harmful, mostly because of the high rates of microbiological contamination, turbidity, iron, fluoride, and aluminium.

References:

[1] A. Alamgir, M. A. Khan, J. Schilling, S. S. Shaukat, and S. Shahab, "Assessment of groundwater quality in the coastal area of Sindh province, Pakistan," Environmental Monitoring and Assessment, vol. 188, no. 2, pp. 1-13, 2016.

[2] M. A. Baluch and H. N. Hashmi, "Investigating the impact of anthropogenic and natural sources of pollution on quality of water in Upper Indus Basin (UIB) by using multivariate statistical analysis," Journal of Chemistry, vol. 2019, 2019.

[3] F. Misaghi, F. Delgosha, M. Razzaghmanesh, and B. Myers, "Introducing a water quality index for assessing water for irrigation purposes: A case study of the Ghezel Ozan River," Science of the Total Environment, vol. 589, pp. 107-116, 2017.

[4] M. Kavurmaci and A. K. Üstün, "Assessment of groundwater quality using DEA and AHP: a case study in the Sereflikochisar region in Turkey," Environmental monitoring and assessment, vol. 188, no. 4, pp. 1-13, 2016.

[5] I. Khan et al., "Impact of city effluents on water quality of Indus River: assessment of temporal and spatial variations in the southern region of Khyber Pakhtunkhwa, Pakistan," Environmental monitoring and assessment, vol. 190, no. 5, pp. 1-19, 2018.

[6] C. Maria Cavadas Morais Couto et al., "Spatiotemporal distribution and sources of trace elements in Ave river (Portugal) lower basin: Estuarine water, sediments and indigenous flora," International journal of environmental research, vol. 13, no. 2, pp. 303-318, 2019.

[7] S. O. Wolf, "China-Pakistan Economic Corridor and its impact on regionalisation in South Asia," in Regional Cooperation in South Asia: Springer, 2017, pp. 99-112.

[8] A. Hussain, K. Ullah, U. Perwez, and M. Shahid, "The long-term forecast of Gilgit baltistan (GB)'s electricity demand," in 2018 International Conference on Power Generation Systems and Renewable Energy Technologies (PGSRET), 2018: IEEE, pp. 1-5.

[9] Z. Iqbal, S. Shahid, K. Ahmed, T. Ismail, and N. Nawaz, "Spatial distribution of the trends in precipitation and precipitation extremes in the sub-Himalayan region of Pakistan," Theoretical and Applied Climatology, vol. 137, no. 3, pp. 2755-2769, 2019/08/01 2019, doi: 10.1007/s00704-019-02773-4.

[10] M. Ahsan, H. Rasheed, M. Ashraf, and K. Anwaar,, "Assessment of Water Quality Status in Gilgit-Baltistan " p. 42. , 2021.