Proceedings of the 9th International Conference of Recent Trends in Environmental Science and Engineering (RTESE 2024)
July, 2025 | Imperial College London Conference Center, London, United Kingdom
Paper No. 134 (The number assigned by the OpenConf System)
DOI: 10.11159/rtese25.134

Identifying the spatial patterns of water quality indicators of the Lower Mahoning River, OH, U.S.

Sahar Ehsani[1*], Felicia Armstrong[1], Sheikh Mohd Sheeraz[1], Richard Deschenes[1], Izrar Ahmad[2]
[1]Department of Civil/Environmental and Chemical Engineering, Youngstown State University
One University Plaza. Youngstown, OH. USA. 44555
[2]Department of Geology, Aligarh Muslim University, Aligarh, India. 202001
*sehsani@ysu.edu

Abstract
This study spatio-temporally evaluates the water quality parameters of the lower Mahoning River, OH, U.S.A. for the years 2012-2015 and the effects of precipitation and land use on the water quality. The principal component analysis was performed on water quality parameters TDS, TSS, As, Ba, Cd, Cu, Fe, Pb, Mn, Ni, Se, Zn of the lower Mahoning River. The first three components explained 83.742% of the variance in the dataset. The principal components were transformed into composite scores and extrapolated using the inverse distance weighted technique to analyse the relative water quality in ArcGIS. The composite score inverse distance weighted maps show that from 2012 to 2015, the water quality degraded as it flowed downstream. However, the water quality improved in the last stretch of the river downstream. The lower Mahoning River watershed becomes more urbanized moving downstream until the end of the basin when land use becomes more forested and comparatively less urbanized. The composite score inverse distance weighted trends in water quality could be connected to this change in land use. In the upstream region of the river, TSS (0.64) was highly correlated with rainfall events while TDS (-0.57) showed a highly negative correlation, which is indicative of the fact that rainfall could act as a dual characteristic i.e., contributing to TSS concentration and decrease TDS through dilution. In the downstream region, a strong correlation existed between rainfall and TSS (0.69), Cu (0.8), Pb (0.8), Zn (0.75), Ba (0.72), and Mn (0.65). This correlation can be attributed to the fact that the downstream collects the runoff from the urbanized area of the downstream of river. The research findings suggest the influence of precipitation on the water quality of the lower Mahoning River and the impact of forested and vegetated areas on maintaining water quality.

***Keywords*:** Water Quality, Surface Water Pollution, Heavy Metal Contamination, Inverse Distance Weighted (IDW), Principal Component Analysis (PCA), Composite Scores

1. Introduction

Heavy metals are among the most pervasive and harmful contaminants on the planet. Their presence in water can directly or indirectly harm the environment through biological biomagnification. Dangerous levels of heavy metals can disrupt the biological mechanisms of aquatic organisms, affecting their ability to live, reproduce and behave [1],[2]. Heavy metal concentrations in water increase due to runoff from sources such as industrial waste, construction debris, asphalt, vehicle emissions, agricultural chemicals, and household waste [3],[4]. In areas with a pronounced geochemical background, soil erosion and the diffusion of toxic metals into rivers can lead to basin-wide pollution. [5]. Lakes, rivers, and estuaries are utilized worldwide for diverse purposes, including drinking, farming, manufacturing, transportation, agriculture, recreation, and waste disposal. Factors such as land surface features, runoff volume, and land use influence water quality [6]. This study focuses on heavy metal pollution in the Mahoning River. The Mahoning River separates regions of western Pennsylvania and northeastern Ohio, flowing through various Ohio counties before entering Pennsylvania to merge with the Shenango River, and form Beaver River. The river's basin encompasses roughly 1,130 square miles [7]. Historically, Mahoning River served as a means of transportation and as a raw material in several industrial activities. Iron and steel machinery used the river to cool their furnaces and other machinery, and then dumped the contaminated, scorching water back into the river [8], [9].

The Mahoning River is regarded as one of the most contaminated rivers in America, tainted with heavy metals including Cu, Pb, Zn, Ni, and Cd. The accumulation of these pollutants over time has degraded the aquatic environment, endangered both the native flora and fauna, and posed a potential risk to human health [10]-[12].

This study aims to identify the effects of land use and precipitation on the water quality parameters (TDS, TSS, As, Ba, Cd, Cu, Fe, Pb, Mn, Ni, Se, Zn) of the Mahoning River, years 2012- 2015.

The research hypothesis of the study is that the spatial and temporal distribution of heavy metal concentrations in the Mahoning River varies in the upstream and downstream stretches of the Lower Mahoning River (study period: 2012-2015). Precipitation and landuse is expected to influence the water quality in the Mahoning River.

1.1 Study area

The study area focuses on a highly polluted section of the lower Mahoning River. This portion of the river flows through Leavittsburg, Warren, Niles, Girard, Youngstown, Campbell, Struthers, and Lowellville. It spans an approximate distance of 40-50 miles (Figure 1).

2. Data

2.1 Water quality and heavy metal data

The water quality data were obtained from the United States Environmental Protection Agency (U.S. EPA) for the years 2012 to 2015. The research utilized data from 7 monitoring stations in the mainstream of the Mahoning River (Figure 1).

2.2 Climate data

Daily precipitation data for the time period and region were obtained from the Climatic Data Online (CDO) database of the National Oceanic and Atmospheric Administration (NOAA). The data were retrieved from five precipitation monitoring stations within the Mahoning watershed [13]. Furthermore, the stations are categorized into distinct zones based on regions, encompassing both water quality stations and rainfall gauge stations (Figure 1).

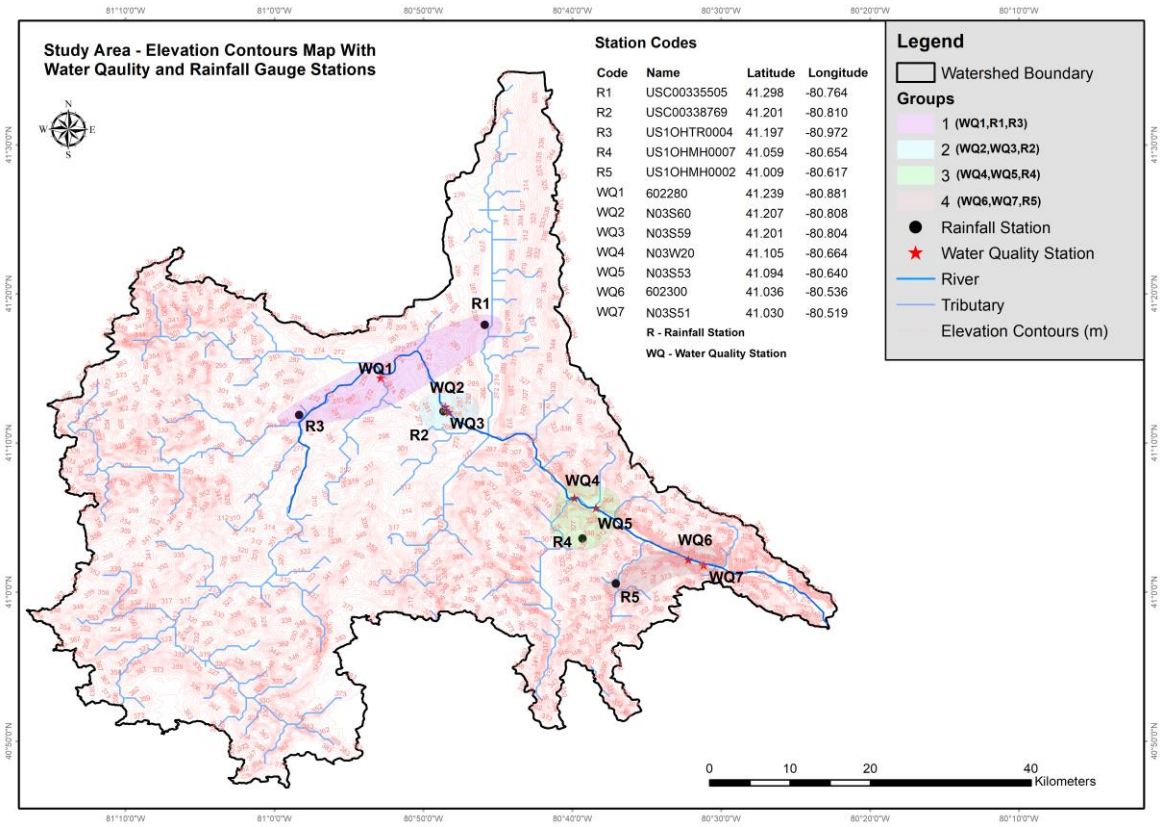

Figure 1 Study Area - water quality and rainfall gauge station of the study area

2.3 Land cover and land use

During the specified timeframe and within the defined region, land cover and land use data were sourced from the United States Geological Survey [14]. Table 1, classifying land cover types, aids in understanding the various land cover categories [15]. The land cover and land use map of the study area is shown in Figure 2. Due to a relatively small change in land use percentage over the study period, one land use map was used for demonstration. The percentage of land use throughout the study period is included in Table 2.

Table 1. Landcover dataset classification.

Value	Characteristic Name	Description
4	Deciduous Broadleaf Forests	The area is characterized by the prevalence of deciduous broadleaf trees, with a canopy height exceeding 2 meters. The percentage of tree cover is greater than 60%.
5	Mixed Forests	The forest area exhibits a relatively equal distribution of deciduous and evergreen tree types, with neither type dominating the canopy which measures over 2 meters in height. The tree canopy coverage exceeds 60%.
8	Woody Savannas	The area under consideration exhibits a tree cover ranging from 30% to 60%, with a canopy height exceeding 2 meters.
9	Savannas	The area exhibits a tree cover ranging from 10% to 30%, with a canopy exceeding 2 meters in height.
10	Grasslands	The area is characterized by the prevalence of herbaceous annuals that do not exceed 2 meters in height.
11	Permanent Wetlands	The lands that are permanently submerged with water covering 30-60% and have a vegetated cover of more than 10%.

12	Croplands	A minimum of 60% of the given region comprises of cultivated cropland.
13	Urban and Built-up Lands	The impervious surface area, comprising building materials, asphalt, and vehicles, should constitute a minimum of 30%.
14	Cropland / Natural Vegetation Mosaics	The cultivation patterns observed in the area under consideration entail mosaics of small-scale farming practices, with a predominant coverage ranging from 40% to 60%. These farming practices are often accompanied by the presence of natural tree, shrub, or herbaceous vegetation.
17	Water Bodies	A minimum of 60% of the region is occupied by bodies of water that are permanent in nature.

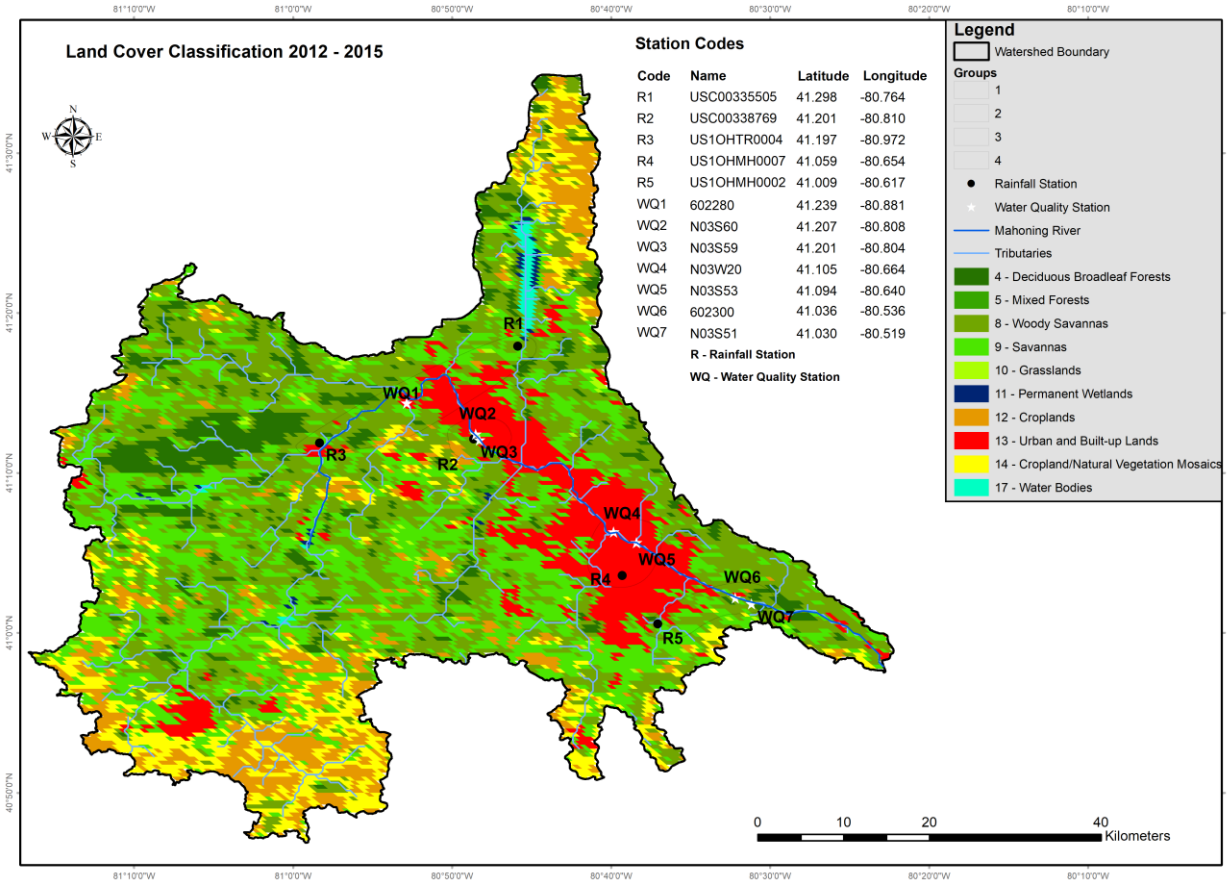

Figure 2 Land use and Land cover maps 2012-15. Land cover and land use data were sourced from the United States Geological Survey [14].

Table 2 Land Cover Classification and Area Percentage Distribution from 2012 to 2015 for the Mahoning River watershed.

Class	Land Cover	Area Percentage (%)					
		2012	2013	2014	2015	Mean in %	Std Dev in %
4	Deciduous Broadleaf Forests	7.75	7.98	7.93	7.75	7.85	0.10
5	Mixed Forests	0.10	0.11	0.10	0.11	0.11	0.01
8	Woody Savannas	34.38	34.09	34.06	34.6	34.28	0.21
9	Savannas	27.14	27.62	27.89	28.72	27.84	0.71
10	Grasslands	0.75	0.85	0.93	0.69	0.81	0.10
11	Permanent Wetlands	0.29	0.27	0.29	0.38	0.31	0.05
12	Croplands	9.48	9.79	9.74	8.56	9.39	0.59
13	Urban and Built-up Lands	10.97	10.97	10.97	10.97	10.97	0.00
14	Cropland/Natural Vegetation Mosaics	8.57	7.74	7.52	7.64	7.87	0.43
17	Water Bodies	0.58	0.57	0.57	0.57	0.57	0.01

3. Methods

3.1 Principal Component Analysis

PCA was used as a dimension-reduction tool to analyse the parameters affecting the chemistry of the Mahoning River [16]. Prior to pre-processing the data for PCA, the suitability of the data is checked using the Kaiser–Meyer–Olkin (KMO) and Bartlett's test of sphericity. PCA applies only when the KMO value is greater than 0.5 and Bartlett's value is less than 0.05 [17]. If the data is found suitable to apply PCA, it is then processed in several steps mentioned below [18].

1. Form a data matrix

$$Y = (y_{ij})_{n*w} = \begin{bmatrix} y_{11} & \cdots & y_{1p} \\ \vdots & \vdots & \vdots \\ y_{n1} & \cdots & y_{np} \end{bmatrix} \quad (1)$$

where y_{ij} is the original observed value, n represents the number of observations and w represents water quality parameters.

2. Standardize the original data using Z-score standardization
3.

$$y_{ij}^{*} = \frac{(y_{ij} - \bar{y}_j)}{SD_j} \quad (2)$$

where $\bar{y}_j$ is the mean value for the jth indicator, y_{ij}* is the standard variable, and SD_j is the standard deviation for the jth indication.

4. Calculate the correlation coefficient matrix, R with y_{ij}^{*}among various parameters
5.

$$R = (r_{ij})_{w*w} = \frac{1}{n-1} \sum_{i=1}^{n} y_{ij}^{*} * \bar{y}_j \quad (3)$$

where, i,j = 1,2,…,w.

6. Calculate eigenvalue (λ) and eigenvector (v) for R, which are then used to calculate the principal components. The eigenvalues account for the variance and contribution of the components and the Principal Components should account for more than 80 % variance in the data.

All the mentioned calculations were performed using the SPSS-22 software.

3.2 Composite Scores

The principal components, linear combinations of the original variables, capture data variance. Scores obtained by projecting variables onto principal components are called factor scores, or component scores. A composite score is

a single summated score calculated based on the variance of each principal component and provides a comprehensive assessment of the overall health of a water body by combining several factors or components into one score that can then replace the original set of components. When looking for differences of water quality in different regions, the composite scores can be analyzed instead of individual variables or principal components. The composite score is calculated using the following formula:

$$F = \frac{\lambda_1}{\lambda_1 + \lambda_2 + \ldots + \lambda_n} F_1 + \frac{\lambda_2}{\lambda_1 + \lambda_2 + \ldots + \lambda_n} F_2 + \cdots \frac{\lambda_n}{\lambda_1 + \lambda_2 + \ldots + \lambda_n} F_n \quad (4)$$

where λ: variance of principal component, n is the number of principal components, F1, F2..Fn is Factor scores [17].

3.3 Inverse Distance Weighting

Inverse Distance Weighting (IDW) interpolation, a widespread spatial analysis method utilizes a linear-weight combinations of available sample spaces to calculate unknown values. It operates on the principle that values of unknown points are more heavily influenced by nearby known points [18]. The IDW operations was performed using ArcGIS software.

4. Results and Discussion

4.1 Principal Component Analysis

The purpose of PCA was to sift through the large data set to identify the primary information that is distinctive of the Mahoning River aquatic environment. The computed results showed that KMO = 0.663 (greater than 0.5), and Bartlett's test of sphericity value = 0 (less than 0.05), which indicated that the data were suitable for PCA. Upon performing the PCA three principal components, PC1, PC2, and PC3, were identified with eigenvalues of 7.396, 4.165, and 1.838 respectively. These eigenvalues corresponded to variances of 46.22%, 26.03%, and 11.48% respectively, accounting for a cumulative variance of 83.74%.

4.2 Factor Score and Composite Score

Since the composite score is negatively correlated with the river water quality, smaller composite score values should indicate better water quality [18]-[20].

4.3 Spatial and Temporal Distribution of Composite Score

The yearly composite scores were calculated to gain insights into the fluctuations observed in water quality stations (Figure 3). The composite score plots indicated that the general characteristics of the spatial distribution of scores remained the same throughout the four years, 2012 to 2015. The composite scores increased from upstream to downstream of the Mahoning River. The runoff, which carries various pollutants from surrounding areas, can contribute to the degradation of water quality as the river continues its downstream flow. Another contributing factor to the deterioration of water quality is the increasing intensity of urban and built-up areas in the downstream sections of the river which can be evident in Landcover classification map (Figure 2, Table 2). In urbanized regions, anthropogenic actions such as industrialization, urbanization, and transportation could discharge contaminants into the aquatic environment, comprising of heavy metals, fertilizers, pesticides, and other chemical substances. The cumulative effect of these intensified urban areas along the downstream stretches of the river further exacerbates the deterioration of water quality. Stations WQ6 and WQ7, situated downstream, consistently exhibited moderate water quality. Being located downstream typically subjects these stations to potential pollutants. However, the surrounding dense forests and natural vegetation act as a natural buffer, capturing these pollutants and minimizing surface runoff, which can carry contaminants into water bodies. This forested environment, combined with limited urbanization in the vicinity, counteracts the potential decline in water quality often seen in downstream locations. Another reason might be the possibility of groundwater from adjoining aquifer contributing to mixing and dilution phenomena locally, highlighting the role of natural environmental determinants in these specific locales.

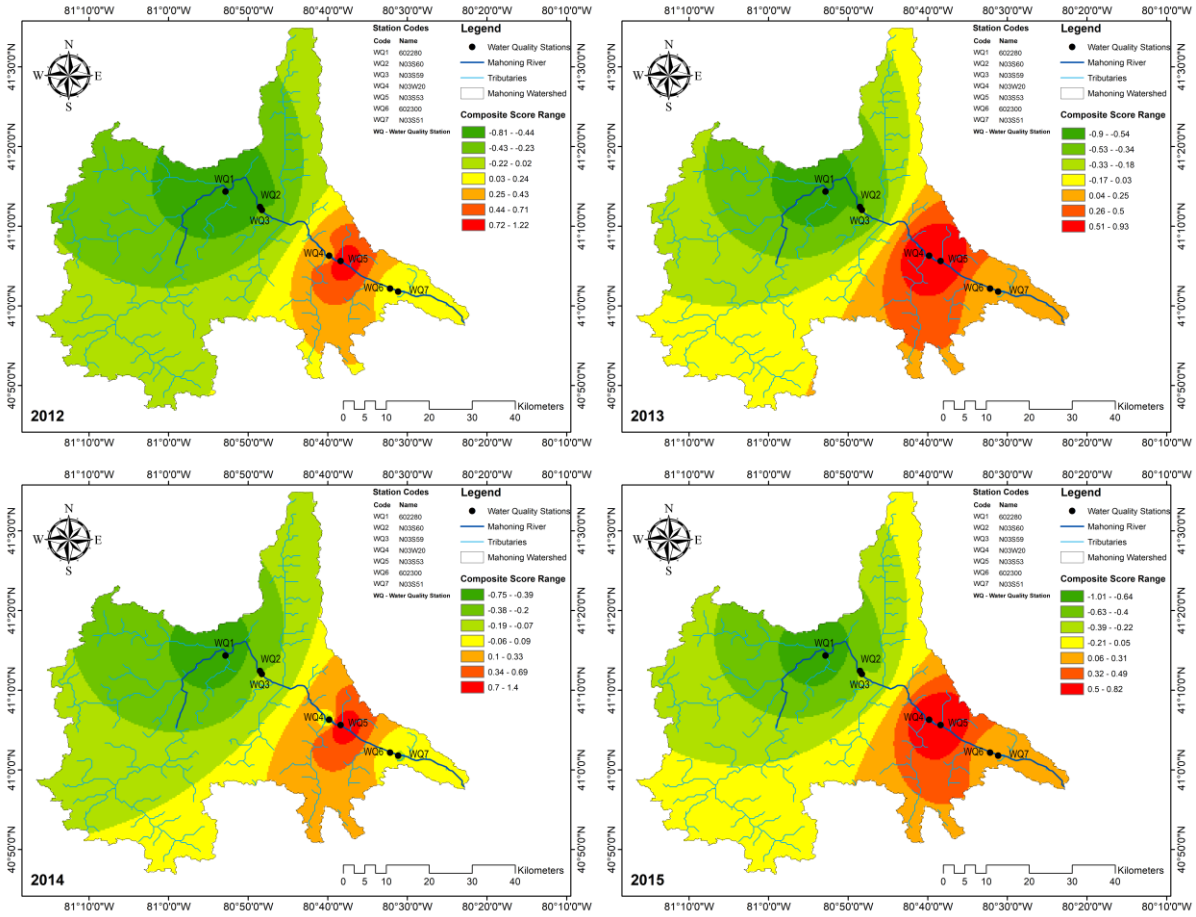

Figure 3 Spatial and temporal distribution of composite scores from 2012 to 2015.

3.6 Precipitation impact on the water quality of the Mahoning River

The impact of rainfall events on water quality data was examined for different Groups (Figure 1). Table 3 presents statistical correlations between selected rainfall events and the water quality parameters. The studied rainfall events had more than 2.5 cm (1 inch) of continuous rainfall and water quality parameters selected were within 24 hours following each rainfall event.

Table 4 Group wise correlation (*r* value) between rainfall events >2.5 cm and various parameters

Parameters	Rainfall events		
	Group 1	Group 2	Group 4
TDS	**-0.57**	0.15	0.43
TSS	**0.64**	0.20	**0.69**
Zn	0.52	0.35	**0.75**
As	0.07	0.27	-0.02
Ba	-0.35	-0.36	**0.72**
Cd	0.00	0.51	0.36
Cu	**0.64**	**0.63**	**0.80**
Fe	0.48	0.27	0.46
Pb	0.52	0.30	**0.80**
Mn	-0.15	0.40	**0.65**
Ni	-0.10	0.27	0.39
Se	0.21	0.42	0.00

Group-wise comparisons (Figure 1) provided insights into the contribution of precipitation to pollution in the Mahoning River. In group 1, TSS and Cu were correlated with the rainfall events (*r* values of 0.64 and 0.64). The TSS

includes sediment, silt, sand, plankton, algae, and other solid impurities such as chemical precipitates. The correlation of TSS with precipitation may be due to the fact that rainfall events potentially contribute to increased levels of suspended solids in the river through urban runoff. Interestingly, TDS showed a highly negative correlation with rainfall events in Group 1. This is likely due to the dilution effect where increased rainfall led to a decrease in the concentration of dissolved solids in the water. (IPCS, 1998). Thus, a dual and contrasting characteristic of rainfall events is noteworthy manifested as elevated concentration of TSS and decreased concentrations in TDS.

The high correlation between Cu and rainfall in all groups may suggest a land-use driven genesis. Cu is released into water through weathering of soil, industrial discharge and urban runoff. The correlation in group 3 could not be drawn due to paucity of precipitation data. In group 4, a strong correlation existed between rainfall events and TSS, Cu, Pb, Zn, Ba and Mn. Whilst TSS is indicative of recent pollution loads, the rest of the parameters such as Pb, Zn, Ba and Mn infer pollution load contributions from multiple sources. Another possible reason of the mentioned correlation is the location i.e., in the downstream direction. The extreme downstream area receives both urban and rural runoff from the upstream and the rest of the sub-catchments hence bears the signature in the form of elevated concentrations of TSS, Cu, Pb, Zn, Ba and Mn.

Additionally, zinc (Zn) and lead (Pb) both showed a correlation of 0.6 or above with rainfall events in two of the groups, indicating a significant impact from runoff processes. Zinc is commonly utilized in industrial processes and is also present in fertilizers and pesticides, representing some of the prevailing sources within the Mahoning River catchment area. Similarly, lead, often associated with industrial processes and the decay of infrastructure such as old pipes and paint, showed a moderate to high correlation with rainfall in two groups, suggesting that runoff from these sources may be contributing to its presence in the river.

There was a weak correlation between As, Cd and Ni and rainfall events. This suggests that the presence of pollution, from these contaminants might be attributed to historical practices than being influenced by recent runoff.

5. Conclusion

This study analyses the spatial distribution and temporal patterns of the water quality parameters of the lower Mahoning River in the years 2012-2015 and examines the effects of land use patterns and precipitation on the river's quality. The strong correlations between rainfall and TSS (0.69), Cu (0.80) in the upstream region, Cu (0.80) in the middle reaches and TSS (0.69), Cu (0.8), Pb (0.8), Zn (0.75), Ba (0.72) and Mn (0.65) in the downstream region underscore the impact of precipitation-driven pollution on the river's water quality.

Integrating land use and landcover patterns with composite scores lend crucial information about water quality of the Mahoning River. In urbanized or industrialized regions, the surface runoffs as a result of precipitation events can transport contaminants into the river, thereby causing elevated concentrations of the contaminants. The assessment of various monitoring stations, in terms of composite scores, indicates that the water quality of the stations in the upstream interface of urban and forested areas, such as WQ1(-0.874), WQ2(-0.778), and WQ3(-0.234), contrast with those in the downstream areas within urban and built-up regions, like WQ5(0.695) and WQ4(0.699). Stations predominantly nestled in forested surroundings, such as WQ7(0.218) and WQ6(0.274), showcased moderate water quality. The findings of the study highlight the interconnectedness of land use patterns and precipitation on the water quality of the lower Mahoning River.

References

[1] S. Mitra, P. Chakraborty, N. Paul, D. K. Mukherjee, R. Das, and S. Gupta, "Impact of heavy metals on the environment and human health: Novel therapeutic insights to counter the toxicity," *Journal of King Saud University - Science*, vol. 34, no. 3, Art. no. 101865, 2022, doi: 10.1016/j.jksus.2022.101865.

[2] S. Ehsani, D. James, and Z. M. Oskouie, "Determining selenium speciation by graphite furnace atomic absorption spectrometry," *Environmental Monitoring and Assessment*, vol. 193, pp. 1–12, 2021.

[3] National Oceanic and Atmospheric Administration (NOAA), "About our agency," 2023. [Online]. Available: https://www.noaa.gov/about-our-agency

[4] G. Yu, F. Chen, H. Zhang, and Z. Wang, "Pollution and health risk assessment of heavy metals in soils of Guizhou, China," *Ecosystem Health and Sustainability*, vol. 7, no. 1, 2021, doi: 10.1080/20964129.2020.1859948.

[5] T.-T. Lim, J.-H. Tay, and C.-I. Teh, "Significance of aqueous cation composition on heavy metal mobility in a natural clay," *Water Environment Research*, vol. 74, no. 4, pp. 346–353, 2002, doi: 10.2175/106143002x140107.

[6] M. M. Bahar, H. Ohmori, and M. Yamamuro, "Relationship between river water quality and land use in a small river basin running through the urbanizing area of Central Japan," *Limnology*, vol. 9, no. 1, pp. 19–26, 2008, doi: 10.1007/s10201-007-0227-z.

[7] A. B. Sullivan, G. M. Georgetson, C. E. Urbanczyk, G. W. Gordon, S. A. Wherry, and W. B. Long, "Modeling flow and water quality in reservoir and river reaches of the Mahoning River Basin, Ohio," U.S. Geological Survey Scientific Investigations Report, 2023, doi: 10.3133/sir20225125.

[8] Riano, "Field Notes: The Mahoning River as Urban Reinvention," *The Architectural League of New York*, 2021. [Online]. Available: https://archleague.org/article/mahoning-valley-dam-removal/

[9] U.S. Environmental Protection Agency (EPA), "Mahoning River Waste Load Allocation Study," 2021. [Online]. Available: https://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=5000113Q.TXT

[10] M. Jaishankar, T. Tseten, N. Anbalagan, B. B. Mathew, and K. N. Beeregowda, "Toxicity, mechanism and health effects of some heavy metals," *Interdisciplinary Toxicology*, vol. 7, no. 2, pp. 60–72, 2014, doi: 10.2478/intox-2014-0009.

[11] T. Liang, H. Wang, X. Zhang, S. Zhang, and X. Yu, "Transportation processes and rates of heavy metals in an artificial rainstorm runoff under different land use types," *Chinese Journal of Applied Ecology*, vol. 14, no. 10, pp. 1756–1760, 2003.

[12] Z. Wang and G. Lei, "Study on penetration effect of heavy metal migration in different soil types," *IOP Conference Series: Materials Science and Engineering*, vol. 394, no. 5, Art. no. 052033, 2018, doi: 10.1088/1757-899X/394/5/052033.

[13] National Oceanic and Atmospheric Administration (NOAA), "About our agency," 2023. [Online]. Available: https://www.noaa.gov/about-our-agency

[14] U.S. Geological Survey, "LP DAAC - MCD12Q1 v006," 2023. [Online]. Available: https://lpdaac.usgs.gov/products/mcd12q1v006/

[15] D. Sulla-Menashe and M. A. Friedl, "User Guide to Collection 6 MODIS Land Cover Dynamics (MCD12Q2) Product," User Guide, vol. 6, no. 1, pp. 1–8, 2018.

[16] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 5, no. 2, pp. 125–143, 2013, doi: 10.1002/wics.1246.

[17] IBM, "KMO and Bartlett's Test of Sphericity," IBM Documentation, 2025. [Online]. Available: https://www.ibm.com/docs/en/spss-statistics/25.0.0?topic=detection-kmo-bartletts-test

[18] W. Yang, L. Zhang, X. Lu, and T. Xue, "Using principal components analysis and IDW interpolation to determine spatial and temporal changes of surface water quality of Xin'anjiang River in Huangshan, China," *International Journal of Environmental Research and Public Health*, vol. 17, no. 8, Art. no. 2942, 2020, doi: 10.3390/ijerph17082942.

[19] F. Wu, Z. Zhuang, H. Liu, and Y. Shiau, "Evaluation of water resources carrying capacity using principal component analysis: An empirical study in Huai'an, Jiangsu, China," *Water*, vol. 13, no. 18, Art. no. 2587, 2021, doi: 10.3390/w13182587.

[20] S. Xu, L. Ji, Y. Chen, Y. Zhao, Y. Yang, and Q. Wang, "The fuzzy comprehensive evaluation (FCE) and the principal component analysis (PCA) model simulation and its applications in water quality

assessment of Nansi Lake basin, China," *Environmental Engineering Research*, vol. 26, no. 2, Art. no. 200022, 2020, doi: 10.4491/eer.2020.022.