Complex Linkage between Watershed Attributes and Surface Water Quality: Gaining Insight via Path Analysis

Afed Ullah Khan 1, 3*, Hafiz Ur Rahman 1, Liaqat Ali 2, Muhammad Ijaz Khan 1, Humayun Mehmood Khan 1, Afed Ullah Khan 1, Fayaz Ahmad Khan 1, 3, Jehanzeb Khan 4, Liaqat Ali Shah 1, Kashif Haleem 1, Asim Abbas 1, Izaz Ahmad 1

1 Department of Civil Engineering, University of Engineering and Technology Peshawar (Bannu Campus), 28100, Bannu, Pakistan.
2 Department of Sociology, University of Swabi, Swabi, 23430, Pakistan.
3 National Institute of Urban Infrastructure Planning, University of Engineering and Technology Peshawar, 25000, Pakistan.
4 Government Post Graduate College Kohat, Higher Education Department KPK, Pakistan.

Received 05 January 2021; Revised 23 March 2021; Accepted 30 March 2021; Published 01 April 2021

Abstract

Understanding the influence of various variables on surface water quality is extremely important for protecting ecosystem health. The principal aim of this study is to assess the direct (DE), indirect (IE) and total effects (TE) of socio-economic, terrestrial and hydrological factors on surface water quality via path analysis through the lens of 15 sub-basins located on Indus basin, Pakistan. Four path models were selected based on Comparative Fit Index (CFI) = 0.999 value. First path model showed that rangelands having low population density decline river runoff which decreases instream Electrical Conductivity (EC) because of lower anthropogenic activities. Second path model depicted that croplands having higher population density enhance river runoff due to irrigation tail water discharge which decline instream EC because of dilution. Third path model showed that croplands with higher population density enhance river runoff which increases instream NO3− concentration because of unscientific application of irrigation water. Fourth path model unveiled that croplands enhance Gross Domestic Product (GDP) which enhance river runoff and instream NO3− concentration. To protect ecosystem health, Best Management Practices (BMPs), precision farming and modern irrigation techniques should be adopted to reduce irrigation tail water discharges containing pollutants entry in Indus River.

Keywords: Path Analysis; Water Quality; Socio-economic; Terrestrial; Hydrological.

1. Introduction

Catchment hydrology is mainly influenced by watershed topography which includes catchment area, shape, and slope [1–4]. Study on Xiangxi River revealed that 26% instream water quality variations was caused by topographic features [4]. Watershed slope has positive relationship with nutrients (Total Phosphorous (TP) and Total Nitrogen (TN)), suspended solid, Chemical Oxygen Demand (COD) and Biological Oxygen Demand (BOD) [5]. Elevation and mean slope have a positive linkage with dissolved oxygen (DO) while negative correlation with dissolved phosphorus, turbidity, EC, COD, water temperature, TN, TP and NO3−N [4, 6]. Some studies reported that standard deviation of

*Corresponding author: afedullah@uetpeshawar.edu.pk
http://dx.doi.org/10.28991/cej-2021-03091683

© 2021 by the authors. Licensee C.E.J, Tehran, Iran. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).
slope enhance the contaminant concentration in surface waters [7, 8]. Literature shows that elevation and mean slope speed up soil erosion which pollute surface waters [4, 9].

Dynamics of land use, socioeconomic activities and fertilizer application causes spatiotemporal variation in surface water quality [10]. Fertilization and irrigation have strong consequences on water-land use relationship [11]. Under and over fertilization have strong effects on surface water quality and soil nutrition [35]. Water quality deterioration (nutrients and fecal indicator bacteria) is linked with animal grazing which threatens human health and the environment [12]. Population density, animal farming and dissolved phosphorous has high influence on GDP, EC, TP, and turbidity [3, 13]. Total suspended solids, NO$_3$+NO$_2$, Cu, Zn, oil and grease have strong linkage with human population density [3, 14, 15]. Intense anthropogenic activities and fertilizer applications for high crop yield badly impacts river water quality [16]. Therefore, it will be significant to reveal the possible complex terrestrial, socioeconomic and hydrologic impacts on surface water quality.

Surficial geology in addition to anthropogenic biomes is primary factors which should be considered in land management policies formulation [17]. Land use-surface water quality relationship has been extensively evaluated by researchers [18, 19, 20]. Intense agricultural activities and urban sprawl has severely affected the aquatic ecosystem health i.e., high overland flow, considerable trace elements and nutrients loads [3, 21]. These relationships suggest the unavailability of safe water for human consumption in near future [4, 9].

Precipitation and irrigation tail water discharges fuel the problem of water quality impairment from nonpoint sources of croplands [19]. NH$_4$-F and NO$_X$-F have strong correlation with cropland biomes [22]. Literature shows that agriculture lands enhance nutrients concentration in surface water bodies which is mainly attributed to irrigation tail water discharges and precipitation overland flow which sweeps all kinds of pollutants especially fertilizer remaining’s from the top fertile soil layer to nearby water bodies [5, 23].

Rangelands play a key role in controlling nonpoint source pollution. Dissolved organic carbon and NO$_X$-F are negatively associated with rangeland biomes [24]. Rangeland’s biomes have lower human interference as well as good nutrients retention capacity which help in water environment protection [25-27]. Degradation of rangeland leads to degradation of water environment [28, 29].

Literature shows that researchers linked socioeconomic, topographic, terrestrial, and hydrological determinants with surface water quality separately [30]. The novelty of the current study is that it simultaneously assesses the DE, IE and TE of the aforementioned determinants on surface water quality. The major focus of this study is to evaluate the complex DE, IE and TE of terrestrial, socioeconomic and hydrologic variables on surface water quality in Indus basin using path analysis approach. Generally, path analysis is used to assess anthropogenic DE and IE on instream water quality. Furthermore, the notion of the current hypothetical model is completely based on literature that how does watershed socioeconomic, terrestrial and stream characteristics interact with each other?

2. Study Area, Data Collection and Methods

2.1. Study Area

Indus river basin is stretched over four countries with total covered area of 1.12 million Km$^2$ in which Pakistan, India, China, and Afghanistan covers 47, 39, 8 and 6% area. Indus river basin spreads over 520000 km$^2$ which covers 65 % territory of Pakistan. Annual precipitation varies geographically from 100 to 500 mm in low elevated areas and 2000 mm in high elevated areas. The flow regime is mainly governed by snowfall at higher altitudes. Indus basin climate varies greatly.

The present study is based on Indus river basin, Pakistan. This study covers fifteen water quality monitoring sites which include Barasin, Draband, Bisham Qila, Bunji, Dadu Moro Bridge, Gunji Bridge, Kachura, Khairabad, Kharmong, Mandori, Massan, Partab Bridge, Rikot, Shatial Bridge, and Sehwan as demonstrated by Figure 1.
2.2. Data Collection

2.2.1. Water Quality and Hydrological Data

Both discharge and water quality data were obtained from Water and Power Department (WAPDA) Pakistan for a period of (1963-2009). The current study covers fifteen monitoring stations which include Barasin, Draband, Bisham Qila, Bunji, Dadu Moro Bridge, Gunji Bridge, Kachura, Khairabad, Kharmong, Mandori, Massan, Partab Bridge, Rikot, Shatial Bridge, and Sehwan. This study is based on twenty-fundamental water quality variables which includes Cl, HCO$_3$, Ca, Mg, Na, K, SO$_4$, NO$_3$, CO$_3$, F, Total Cations and Anions, SiO$_2$, Fe, B, D.S by Evaporation, EC $\times 10^6$ at 25°C, pH, Residual Carbonate me/l, and Sodium Adsorption Ratio (SAR)).

2.2.2. Socioeconomic Data

Digital Elevation Model (DEM) data, downloaded from Shuttle Radar Topography Mission (SRTM), was used for delineating sub-watersheds. For the delineated sub-watersheds, population density data was extracted from Socioeconomic Data and Applications Center (SEDAC) [31]. GDP data was obtained from the provincial economies report [32]. Moreover, provincial level GDP data were assigned to the delineated sub-watersheds.

2.2.3. Terrestrial Determinants Data

Anthropogenic biomes are classified in six main groups which are demonstrated by (Figure 2). Socioeconomic Data and Applications Center (SEDAC) source was used for downloading anthropogenic biomes data. In addition, surficial geology data that includes silt, sand, clay, and gravel was collected from the World Harmonized Soil Database (HWSD).
2.3. Methods

2.3.1. Watershed Modeling

The Digital Elevation Model (DEM) was utilized for delineating sub-watersheds using monitoring points as outlet via spatial analysis tool of Geographical Information System (GIS). DEM data was extracted for the study area as obvious from the (Figure 3). Land use, soil and population data was extracted using GIS. The extracted variables were linked with water quality parameters using path analysis technique. Flowchart demonstrating the methodology of the research is obvious from Figure 5.
2.3.2. Statistical Modeling

Structural equation modeling (SEM) is a multivariate statistical framework which was used to compute direct and indirect complex linkages among social variables, terrestrial variables, river discharge and surface water quality parameters. A hypothetical model of interrelationships among social variables, terrestrial variables, hydrological and water quality parameters is demonstrated by Figure 4.

2.3.3. Data Analysis

The current study is based on mean values of each variable of the conceptual hypothesized path model. Path analysis was carried out to model relationships (DE and IE) among social, terrestrial, hydrological and water quality parameters. The hypothesized model based on four attributes groups was selected as per the available literature. Variables with the probability of strong relationship with certain variables within each of the four groups were also selected based on the previous studies. Path analysis was carried out on a variety of hypothetical models to examine the relationship among various combinations of attributes variables and to choose small group of models that best explains linkage between the social, terrestrial, hydrological and water quality variables. Path coefficients (standardized) were derived from path analysis by simultaneously regressing a dependent variable on each independent variable immediately linked to it by an arrow in the hypothetical path model. Path coefficients were used to evaluate a dependent variable's effects (DE, IE and TE). Independent-dependent variable DE is equal to the corresponding path coefficient. Independent variable's IE was calculated by multiplying coefficients in every path from the independent variable across all of the mediating variables to the dependent variable, and then adding values for all indirect paths [33]. An independent variable's cumulative influence on a dependent variable is equivalent to the number of all DE and IE. The path analyses were performed using AMOS Version 19.0, a Structural Equation

Figure 4. Hypothesized direct and indirect linkage among social factors, terrestrial factors, river discharge and surface water quality parameters. Arrows show associations among hypothetical variables

Figure 5. Flowchart demonstrating the methodology
Modeling (SEM) and path analysis program. The performance of various path models was assessed using Comparative Fit Index (CFI). CFI is preferred for small sample size studies. The CFI value ranges from 0 to 1. Path model having CFI > 0.9 is considered good fit [34].

3. Results and Discussion

Statistical assessment of path models showed that all the four path models are characterized by acceptable CFI values. All the four path models have CFI = 0.999 which shows that modeled values are in close agreement with observed values.

3.1. Path Model: EC as Output Variable

Mean population and discharge are the significant determinants which cause variations in stream EC as obvious from (Figure 6). More specifically mean population has strong direct relationship with riverine flow regime. Pollutants including sediments and heavy metals (Zn, Cu, Pb, etc.) that enhance EC may be discharged to Indus River due to anthropogenic activities i.e. surface runoff, ground water discharge, industrial effluents, soil erosion etc. [35-38]. Mean population has negative linkage with rangelands because rangelands are thinly populated areas. Rangelands are negatively associated with riverine flow regime which may be due to high infiltration owing to its pervious surface area [22, 25-27]. Basin having higher ration of rangelands decline pollutants load in surface waters owing to lower elasticity value and higher pollutants retention capability. Conservation of rangelands at watershed scale can help to protect ecosystem health. Riverine flow regime has negative correlation with EC which might be due to dilution effects [22]. Moreover, 2/3 climate models demonstrated that annual average runoff would increase in future [17]. In conclusion the mean population has strongest direct increasing while riverine flow has direct decreasing effect on EC. The total effect of mean discharge on EC is greater than the mean population and rangelands as obvious from Table 1.

### Table 1. Mean population, rangeland, and mean discharge DE, IE and TE on in stream EC

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Mean population</th>
<th>Rangeland</th>
<th>Mean discharge</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DE</td>
<td>IE</td>
<td>TE</td>
</tr>
<tr>
<td>Rangeland</td>
<td>-0.287</td>
<td>-0.287</td>
<td></td>
</tr>
<tr>
<td>Mean discharge</td>
<td>0.674</td>
<td>0.085</td>
<td>0.759</td>
</tr>
<tr>
<td>EC</td>
<td>0.97</td>
<td>-0.511</td>
<td>0.459</td>
</tr>
</tbody>
</table>

Figure 6. Social, terrestrial, and hydrological variables DE and IE influence on surface water EC

Mean population and riverine flow are the significant determinants which cause variations in stream EC as obvious from (Figure 7). More specifically mean population has strong direct relationship with riverine EC. Pollutants including sediments and heavy metals (Zn, Cu, Pb, etc.) that enhance EC may be discharged to Indus River due to anthropogenic activities i.e. surface runoff, ground water discharge, industrial effluents, soil erosion etc. [35-39]. Mean population has positive linkage with croplands because Indus is the largest basin in Pakistan where people live and grow crops. Riverine flow regime has negative linkage with EC which might be due to dilution effects [24]. Moreover, 2/3 climate models demonstrated that annual average runoff would increase in future [17]. Croplands are negatively associated with instream EC which may be due to irrigation tail water discharges which enhance riverine flow, causing dilution effect, owing to unscientific application of irrigation water in the Indus basin [40-42]. In conclusion the mean population has strongest increasing while riverine flow has decreasing effect on EC. The total effect of mean discharge on EC is greater than the mean population and croplands as obvious from (Table 2).
Wu et al (2015) found that path model having EC as outcome variable demonstrated that road density, population density and base flow ratio causes variation in stream EC. The impacts of human population density on EC are positively and negatively mediated by road density and base flow ratio respectively. Results of the current study are supported by the aforementioned research work [14].

Table 2. Mean population, cropland, and mean discharge DE, IE and TE influence on instream surface water EC

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Mean population</th>
<th>Cropland</th>
<th>Mean discharge</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DE</td>
<td>IE</td>
<td>TE</td>
</tr>
<tr>
<td>Cropland</td>
<td>0.37</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>Mean discharge</td>
<td>0.759</td>
<td>0.759</td>
<td></td>
</tr>
<tr>
<td>EC</td>
<td>1.036</td>
<td>-0.578</td>
<td>0.459</td>
</tr>
</tbody>
</table>

Figure 7. Social, terrestrial, and hydrological variables DE and IE influence on surface water conductivity

3.2. Path Model: Total Nitrogen as Outcome Variable

Mean population and river runoff are the significant determinants which cause variations in NO$_3^-$ concentration as obvious from (Figure 8). River runoff has the strongest direct impact on riverine NO$_3^-$ concentration. Mean population has positive linkage with croplands because Indus is the largest basin in Pakistan where people live and grow crops. Riverine flow regime has positive linkage with NO$_3^-$ because nitrogen flux substantially increases with precipitation [24]. Pakistan lies in the temperate climate zone, generally arid climate, characterized by cold Winters and hot Summer, where the sensitivity of riverine flow to precipitation and snowmelt is higher which fuels the top soil erosion which enhances riverine NO$_3^-$ concentration. Similarly mean population has increasing effect on riverine NO$_3^-$ which may be due to high anthropogenic activities in Indus basin. Areas with high population density intensely nitrogen loads in riverine waters [5, 23, 28, 43-48] which may be due to intense human activities and low retention capacity in such areas [49]. Intense storms increase overland flow which sweeps domestic sewage and nutrients to nearby water bodies [19]. Croplands are negatively associated with instream NO$_3^-$ concentration which may be due to irrigation tail water discharges which enhance riverine flow, causing dilution effect, owing to unscientific application of irrigation water in the Indus basin [40-42]. The total effects of riverine flow regime on NO$_3^-$ concentration overweighs mean population in Indus basin as obvious from (Table 3). This study indicates that streams in this landscape will obtain more contaminants from various human activities [40-42]. It is utmost important to focus on the possible adverse consequences of high nitrogen rates. Various BMPs such as vegetated buffer strips [50, 51] or nitrogen reducing bioreactors should be adopted in such conditions [51, 52].

Table 3. Mean population, cropland, and mean discharge DE, IE and TE on instream NO$_3^-$

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Mean population</th>
<th>Cropland</th>
<th>Mean discharge</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DE</td>
<td>IE</td>
<td>TE</td>
</tr>
<tr>
<td>Cropland</td>
<td>0.370</td>
<td>0.370</td>
<td></td>
</tr>
<tr>
<td>Mean discharge</td>
<td>0.759</td>
<td>0.759</td>
<td></td>
</tr>
<tr>
<td>NO$_3^-$</td>
<td>0.030</td>
<td>0.223</td>
<td>0.252</td>
</tr>
</tbody>
</table>
GDP and river runoff are the significant determinants which cause variations in NO$_3$ concentration as obvious from (Figure 9). But river runoff has the strongest direct impact on riverine NO$_3$ concentration. GDP has positive linkage with croplands because Pakistan is an agricultural country where GDP mainly depend on agricultural products and Indus is the largest basin in Pakistan where people grow crops. Croplands are positively associated with riverine flow regime and negatively linked with NO$_3$ concentration. The aforementioned association may be due to irrigation tail water discharges which enhance riverine flow, causing dilution effect, owing to unscientific application of irrigation water in the Indus basin [40-42]. Riverine flow regime has positive linkage with NO$_3$ concentration because sensitivity of riverine flow to precipitation is higher in Pakistan which enhances nitrogen flux. Surface runoff sweeps and erodes the top fertile soil layer which increase riverine NO$_3$ concentration [24]. GDP is positively linked with riverine NO$_3$ concentration which may be due to urbanization and agriculture activities in Indus basin. Economic growth enhances urban development and agriculture activities where domestic sewage and irrigation tail water discharges enhance riverine NO$_3$ concentration [22]. This study indicates that streams in this landscape will obtain more contaminants from anthropogenic activities [40-42]. It is utmost important to focus on the possible adverse consequences of high nitrogen rates. Various BMPs such as vegetated buffer strips [50, 51] or nitrogen reducing bioreactors should be adopted in such conditions [51, 52].

Wu et al (2015) found that Path model having total nitrogen as outcome variable demonstrated that population density, riverine flow, % crop land causes variation instream nitrogen concentration. The influence of human population density on total nitrogen is positively mediated by riverine flow [14]. The findings of the aforementioned research work support our study results.

Table 4. GDP, cropland, and mean discharge DE, IE and TE on instream NO$_3$

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>DE</th>
<th>IE</th>
<th>TE</th>
<th>DE</th>
<th>IE</th>
<th>TE</th>
<th>DE</th>
<th>IE</th>
<th>TE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cropland</td>
<td>0.218</td>
<td>0.218</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean discharge</td>
<td>0.264</td>
<td>0.049</td>
<td>0.313</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NO$_3$</td>
<td>0.23</td>
<td>0.072</td>
<td>0.302</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 9. Social, terrestrial, and hydrological variables DE and IE on surface water NO$_3$
4. Conclusion

Indus basin has complex socioeconomic, terrestrial, and hydrological conditions for which it can be difficult to identify all possible relationships and mechanisms which deteriorate surface water quality. Here we assessed the DE, IE and TE of socioeconomic, terrestrial and hydrological variables on stream water quality. Instream water EC increases with croplands while decreases with rangelands. The results demonstrated that watersheds having rangeland biomes pose lower risk to instream water EC impairment as compared to cropland. Similarly, instream water NO3 concentration increases with mean population, GDP, and river flow. The results implied that in stream NO3 concentration is enhanced by anthropogenic activities and river flow. Goodness of fit statistics demonstrated that path models are stronger. Small number of parameters entered the models because of small sample size (15 monitoring sites). The variables were selected based on literature to quantify the principal cause and effect models. Difference in linkage of socioeconomic, terrestrial, and hydrological variables with water quality parameters indicated that different activities impact certain water quality parameter differently which should be addressed while formulating water policies. To control elevated level of NO3 concentration BMPs such as vegetative buffer strips, nitrogen reducing bioreactors should be adopted. Overall, the findings of the current study indicated that multiple adaptation strategies should be adopted to protect stream health.

5. Declarations

5.1. Author Contributions
Conceptionalization, A.U.K.; data collection and processing, H.R., M.I.K., H.M.K. and A.U.K.; analysis and interpretation of the data, A.U.K.; writing original draft preparation, F.A.K. and A.U.K.; writing, review and editing, K.H., L.A., L.A.S., J.K., I.A. and A.A. All authors have read and agreed to the published version of the research article.

5.2. Data Availability Statement
The data presented in this study are available on request from the corresponding author.

5.3. Funding
The authors received no financial support for the research, authorship, and/or publication of this article.

5.4. Conflicts of Interest
The authors declare no conflict of interest.

6. References


