Mohammad Shenavaei Abbasi

POWER LAW SCALING OF STREAM-FLOW IN MOUNTAINOUS BASINS

Master Thesis
Faculty of Engineering and Natural Sciences
Examiners: University Lecturer Chunyan Tang
University Instructor Annina Takala, and
Associate Professor and Research Advisor Omar
I. Abdul-Aziz (West Virginia University, USA)
June 2023

ABSTRACT

Mohammad Shenavaei Abbasi: Power law scaling of streamflow in mountainous basins
Master's Thesis
Tampere University
Master of Science (Technology) Degree Program in Environmental Engineering
June 2023

Floods and droughts are significant threats to communities worldwide. Streamflow is a key predictor of riverine floods and watershed-scale drought events. River discharge is a response of various underlying and interacting climatic, land cover, and hydrological processes in the draining watersheds. Spatiotemporal scaling of streamflow regimes (low through high discharge events) may uncover emerging hydrologic patterns, leading to the development of robust prediction models. The main objective of this thesis is to investigate the scaling property of streamflow using statistical modeling and data analytics with a focus on mountainous areas. The goal is to identify the dominant watershed hydroclimatic and land cover drivers controlling various streamflow regimes, as well as to develop power law-based scaling models to robustly predict streamflow at different time and space scales. The dominant controls of streamflow regimes were identified by using Pearson correlation analysis and principal component analysis. Robust scaling models of various streamflow regimes were then developed by using univariate to multivariate power law functions iteratively.

The study area of this thesis is the Monongahela River Basin of USA. It is a large watershed located mostly in mountainous Appalachian regions of USA with variable land cover characteristics. The watershed passes through three states, and the region's varied topography and land-cover make it a unique place to study. Eleven hydro-climatic variables are collected including different land cover characteristics (six categories) drainage density, imperviousness, area, slope, and precipitation percentiles. A cumulative approach is then implemented on all variables to capture scaling at multiple spatial scales. The spatial correlation of variables is then broken using a bootstrap resampling approach to prevent pseudoreplications. After performing regression analyses, the efficiency and accuracy of the model are assessed by using, respectively, the Nash-Sutcliffe efficiency (NSE) and the root-mean-square error to the standard deviation of observations (RSR) ratio.

The results revealed that drainage area and mean precipitation are the most dominant variables that influence all flow regimes in largely forested mountainous basin. Other hydroclimatic and land cover variables exhibited significant mutual correlations. For example, the most common land cover types, vegetated and agricultural land cover, showed a strong correlation with land area and built-up lands, respectively. The principal component analysis results provide information about model collinearity correlation and significant variables. Previous analyses that provide a reliable scaling model are validated. The power-law model also shows promise in predicting streamflow in an ungauged basin, with an average NSE of 0.90 and RSR of 0.27. As a result, studying streamflow scaling not only reveals details about the mechanisms governing streamflow and how it interacts with other variables, but it also provides useful insights into how to predict using simple power law equations. This study provides information about the mechanisms of streamflow in mountainous watersheds, which can help water resource scientists, engineers, and managers predict streamflow regimes and mitigate floods in similar regions, as well as regulate flood management schemes in different watersheds.

Keywords: Scaling, Streamflow, Hydrology, Modelling, Power Law,

The originality of this thesis has been checked using the Turnitin Originality Check service.

PREFACE

It is with great pleasure and a sense of accomplishment that I present this master thesis, which represents the culmination of years of hard work, dedication, and collaboration. This thesis would not have been possible without the support and guidance of my three supervisors, who have been instrumental in shaping my academic journey.

First and foremost, I would like to express my sincere gratitude to Professor Dr. Omar I. Abdul-Aziz from the West Virginia University, USA for his unwavering support and invaluable insights throughout the course of my research. His depth of knowledge and expertise in the field has been crucial in helping me navigate the complex terrain of my research area.

I would also like to extend my heartfelt thanks to Dr. Tang and Dr. Takala from Tampere University, who have played an integral role in shaping my research project and providing me with invaluable feedback and guidance. Their expertise and commitment to excellence have been a constant source of inspiration and motivation for me.

In addition, I would like to express my heartfelt appreciation to my family and friends for their unwavering support, encouragement, and belief in me throughout this journey. Their support has been the backbone of my success, and I am eternally grateful for their love and encouragement.

In conclusion, I am honored to present this master thesis, and I hope that it will make a meaningful contribution to the field. It has been an incredible journey, and I am grateful to have had the support and guidance of such amazing supervisors, family, and friends.

Tampere, June 2023

Mohammad Shenavaei Abbasi

CONTENTS

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

A	Area (Km2)
DD	Drainage Density (Km-1)
Imp	Imperviousness (%)
P	Precipitation (mm/day)
P_T	Precipitation at percentile T (mm/day)
Q_o	observed discharge (m3/s)
Q_P	Predicted discharge (m3/s)
Q_T	Flow at percentile T (m3/s)
r	Pearson correlation coefficient
S	Mean Slope (%)
σ_o	Standard deviation of observations
AGR	Agricultural land
AIC	Akaike Information Criterion
BUL	Built-up lands
CV	Coefficient of variation
EPA	Environmental Protection Agency
MRLC	Multi-Resolution Land Characteristics
NLCD	National Land Cover Database
NOAA	National Oceanic and Atmospheric Administration
NSE	Nash-Sutcliffe efficiency
OPN	Open land
OVB	Omitted-variable bias
PCA	Principal component analysis
PCBs	Polychlorinated biphenyls
RSR	Root-mean-square error to the standard deviation of observations
USA	United States of America
USGS	United States Geological Survey
VEG	Vegetated land
WAT	Waterbody
WET	Wetlands

1. INTRODUCTION

Water-related catastrophes, also known as hydro hazards, are among the natural disasters that occur the most frequently and pose a threat to both people's lives and the advancement of society and the economy. Flooding is an example of one of these hydro hazards, which can have a significant impact on society. Flood dangers can be reduced by flood planning, which involves the establishment of flood risk management systems.

Streamflow, also known as runoff, is the flow of water in rivers, streams, and other channels that results from precipitation, snowmelt, and groundwater discharge. It is an essential component of the hydrological cycle, as it plays a critical role in transporting water from the land surface to the ocean. Streamflow is a complex process that involves various components, including precipitation, infiltration, evapotranspiration, and storage (Adams & Pagano, 2016).

Understanding of streamflow behavior in a catchment level is essential for making flood management strategies. (Alfieri et al., 2012). Analyzing historic streamflow data, trends and seasonal variations helps in identifying areas and time periods prone to flooding. Also, hydraulic and hydrological models are developed based on streamflow characteristics to simulate how flood waters will flow through a region under different conditions like rain intensities, reservoir levels and drainage capacities (Sharma et al., 2021).

Scaling in hydrology refers to the process of relating the characteristics of a hydrological system (such as streamflow, precipitation, or soil moisture) at one scale to another scale (Blöschl & Sivapalan, 1995). The goal of scaling in hydrology is to understand how the processes that control hydrological systems vary across different spatial and temporal scales. Over the last decade, scholars have paid close attention to scaling behaviors of watershed properties (Özger & Singh, 2013; Shen & Chui, 2021). Streamflow scaling refers to long term statistical dependencies of streamflow with one or several watershed characteristics (e.g., drainage area, precipitation, imperviousness, drainage density) (Tessier et al., 1996). Understanding the streamflow's scaling properties provides information about the internal process dynamics and underlying mechanisms governing how streamflow develops, accumulates, and changes over time based on various hydrological processes and how precipitation translates into runoff. Also, it can be used in long term empirical prediction when planning flood control structures like dams and spillways.

It has been commonly assumed that drainage area is the primary factor influencing streamflow magnitude and temporal scaling (Vogel & Sankarasubramanian, 2000). In this cases streamflow percentiles are represented as a power law function of area (Perez et al., 2018).There have been studies that have used precipitation and catchment descriptors (e.g., area, slope, drainage density, imperviousness, land use) as explanatory factors for streamflow scaling research, as well as flow prediction in gauged and ungauged basins (Abdul Razaq et al., 2016; Formetta et al., 2021; Lacombe et al., 2014; Zheng et al., 2021).

Most research to date has centered around scaling average and peak streamflow regimes as a function of drainage area, precipitation, mean slope and other watershed characteristics. However, streamflow scaling law and behavior of the associated exponents in mountainous large-scale watersheds has been inadequately investigated and fell into the knowledge gap. Here, streamflow scaling features are investigated in mountainous regions, considering the Monongahela River Basin of USA as the study area. On the basis of data collected between 1990 and 2020, the scaling characteristics of the watershed were investigated. The unique feature of this subject is that, compared to the previous scaling studies on streamflow, the current research represents a large (18,959 km^2) mountainous basin, having mostly forest cover alongside notable agricultural and built-up areas. Climatic and watershed drivers are used to investigate the emergent power laws of various streamflow regimes and parsimonious models are estimated to predict streamflow within the watershed. Through this research, fundamental knowledge about hydrologic scaling and organizing principles in watersheds is provided, which may also be used to predict streamflow in ungauged basins depending on climatic, land use, and hydrologic similarities. The specific research questions are as follows:

- What are the main watershed features that control streamflow in mountainous areas with mixed landcovers?
- Does streamflow of various hydrologic regimes present a power law scaling in the mountainous watersheds?
- Can we predict streamflow and examine the hydrological regimes in ungauged areas using scaling properties?

Chapter 2 investigates the background of streamflow, scaling, and streamflow prediction across space and time. Chapter 3 focuses on methodology and the process of data analysis in our case study. The study area and data collection strategies for the analysis is examined. The findings from analyses, scaling exponents, and streamflow prediction are presented in Chapter 4 along with a comparison of the predicted outcomes to actual data. In chapter 5, the results and findings that emerged from our case study is thoroughly examined. The implications and significance of those results were analyzed in depth. In the subsequent chapter 6, the key conclusions that could be drawn from our case study as a whole are summarized. Based on the outcomes that were observed and the insights that were gained, the main conclusions and recommendations that arose from our in-depth analysis of this case is presented.

2. THEORETICAL BACKGROUND

2.1 Streamflow mechanism

There are several factors that affect streamflow (Figure 1). The most important factor influencing streamflow is precipitation. It can occur in various forms, including rain, snow, and hail, depending on the temperature and atmospheric conditions. In mountain areas, snowfall is a significant component of precipitation, and snowmelt plays a crucial role in streamflow generation, particularly during spring and summer months when temperatures rise (Kirkby, 1988). The timing and magnitude of snowmelt are controlled by factors such as snow accumulation, air temperature, solar radiation, and wind. The spatial distribution of snowmelt is also influenced by the topography and vegetation, which can create shading and wind sheltering effects (Barnett et al., 2005).

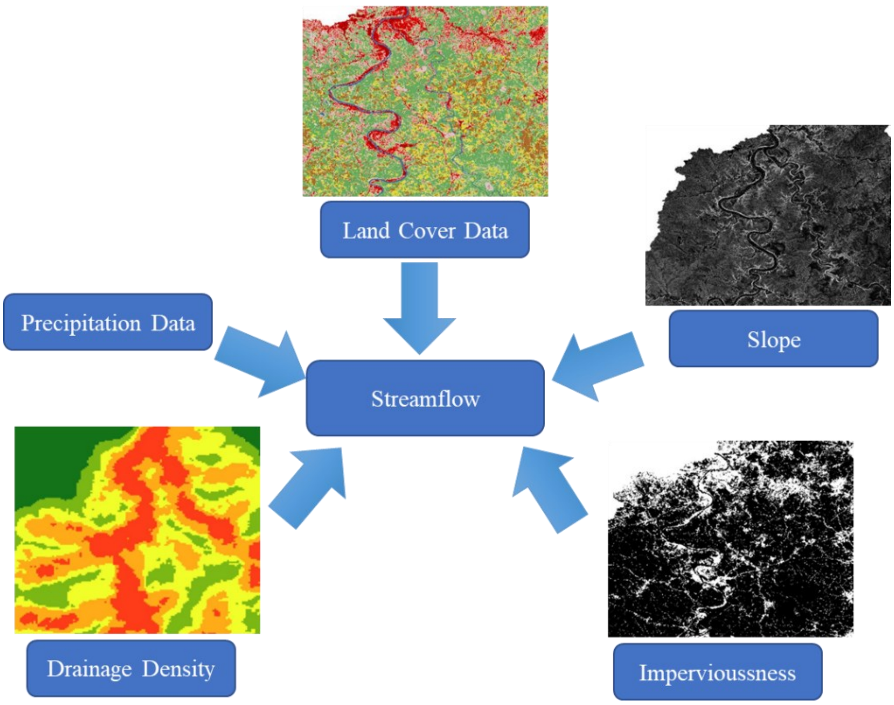

Figure 1. Streamflow mechanism with other climatic and hydrologic variables

Water enters the soil through the process of infiltration from the surface. The rate of infiltration is influenced by factors such as soil texture, soil structure, and vegetation cover. In mountain areas, the soil may be shallow, rocky, or frozen, limiting the infiltration capacity and leading to higher runoff generation (Berghuijs et al., 2014). Soil moisture is a critical component of the streamflow mechanism, as it controls the partitioning between infiltration and runoff. During wet periods, when soil moisture is high, a larger proportion of precipitation will contribute to runoff. Conversely, during dry periods, a larger proportion of precipitation will be absorbed by the soil, reducing runoff generation (Thomas et al., 2021).

The process by which water is transferred from the land surface to the atmosphere via evaporation and plant transpiration is known as evapotranspiration. Evapotranspiration is influenced by factors such as temperature, humidity, wind speed, solar radiation, and vegetation cover. In mountain areas, evapotranspiration rates can vary significantly due to variations in elevation, aspect, and vegetation (Ellenburg et al., 2018).

Interaction of runoff with other variables, including topography, slope, drainage density, and land cover characteristics can hugely impact streamflow (Figure 1). The slope of catchment has a significant influence on streamflow. In general, steeper slopes result in higher flow velocities and increased potential for erosion (Leopold et al., 1964).Steeper slopes also enhance the channelization of runoff, which can lead to higher peak streamflow and decreased infiltration. Moreover, the slope affects the hydraulic geometry of the river, with steeper slopes typically associated with a more narrow and deep channel cross-section (Leopold & Jr., 1953).

Drainage density is a measurement of the total length of streams per unit area and is an essential parameter in the characterization of catchment hydrology (Horton, 1945). High drainage density indicates a well-developed drainage network, which can efficiently transfer water from the landscape to the main channel (Dunne & Leopold, 1958). High drainage density areas usually result in faster response times to precipitation events, increased overland flow, and reduced infiltration. Conversely, low drainage density areas may promote infiltration and groundwater recharge, reducing the contribution of surface runoff to streamflow (Horton, 1945).

Land cover is another crucial factor affecting streamflow. Different types of land cover have varying effects on the hydrological processes of a catchment. For example, forested areas with high canopy interception and evapotranspiration rates can reduce surface runoff and promote infiltration (Bosch & Hewlett, 1982). Conversely, urban or agricultural areas with impervious surfaces or compacted soils can limit infiltration and increase overland flow, leading to higher streamflow volumes and peak discharge rates (Arnold & Gibbons, 1996).

2.2 Scaling theory

Scaling is classified based on the scaling exponent (Gupta, 2004). The relationship between streamflow and a watershed variable, such as drainage area or precipitation, is referred to as simple scaling in streamflow when exponents are constant for all flow quantiles. In other words, the coefficient of variation (CV) of floods remains constant in simple scaling when the flow regimes change. Simple scaling relationships are often described by a power law relationship, in which streamflow is related to variables raised to a power (Gupta & Dawdy, 1995). Multi-scaling, on the other hand, is another scaling type when scaling exponents varies by different flow quantiles. Identifying type of scaling in a study is done after data analyses and calculating coefficient of variation (CV) of floods (Gupta & Waymire, 1990).

The study of scaling properties of rivers and streamflow has a long history in hydrology. Early observations dating back to the mid-20th century noted that the morphology and geometry of river networks exhibit scale-invariant properties (Hack, 1957). In other words, river networks look statistically similar at different scales of observation. This implies there are some universal organizing principles independent of scale. In the 1980s, hydrologists observed that streamflow records across many rivers at different locations exhibit power-law scaling for flood frequencies and magnitudes (Kirkby, 1988). This means that the probability of occurrence of a flood is inversely proportional to its size raised to some power. This scaling exponent was found to be similar across different rivers, suggesting universal behavior.

Further work expanded on the prevalence of scaling in hydrology, including scaling in rainfall (Fraedrich & Larnder, 2016), snowmelt (Deems et al., 2006), and streamflow (Buttle & Eimers, 2009; Tessier et al., 1996). After decades of research, scaling has emerged as a fundamental property in hydrology and fluvial systems with important implications for modeling, monitoring, and ecosystem dynamics.

There are different types of scaling in hydrology, including temporal scaling, spatial scaling, and process scaling (Gupta & Dawdy, 1995). Temporal scaling is the study of the evolution of hydrological variables over time and the relationship between these evolutions and various time scales, such as daily, monthly, or annual. For instance, the examination of streamflow variations through time and their connections with precipitation and temperature (Villarini et al., 2011). Spatial scaling refers to the study of how hydrological variables change across different spatial scales, such as from point to catchment scale. For example, the study of how streamflow changes across different drainage areas, and how these changes are related to precipitation and land use (Medhi & Tripathi, 2015;

Perez et al., 2018). Process scaling refers to the study of how the physical and biological processes that control hydrological systems vary across different scales, such as the study of how evapotranspiration changes across different scales and how these changes are related to vegetation and soil moisture (Ayalew et al., 2014; Thomas et al., 2021).

There has been much research on scaling properties in river flow and streamflow. Pioneering studies (Gupta & Dawdy, 1995; Gupta & Waymire, 1990) established the concept of scaling and an original framework for investigating spatial or regional patterns of flood frequency and the underlying physical mechanisms. In a power law relationship between flow quantiles and drainage area, they investigated scaling exponents for drainage area. Vogel et al (2000) attempted to determine the spatial scaling features of annual streamflow and drainage area in all regions of the United States. The exponents of simple scaling models had a perfect linear correlation with moment order for all regions of the United States. However, it did not produce regional homogeneity of streamflow since scale is insufficient to explain annual streamflow in more various climatic regions (Vogel et al., 1999; Vogel & Sankarasubramanian, 2000). The spatial scaling properties of average annual streamflow across Canada were examined by Yu and Gan (2004), who discovered that a simple scaling law exists between streamflow and drainage area. Furthermore, across Canada and for each of its sub-climatic regions, there was a logarithmic relationship between probability weighted moments and drainage area (Yue & Yew Gan, 2004). Sauquet et al. (2008) conducted a study on 34 French discharge gauging stations to investigate the scaling properties of streamflow using three different statistical methods. These methods included shape analysis of flood hydrographs, analysis of multifractal properties, and the analysis of flood duration frequency curves (Sauquet et al., 2008). They examined the connection between scaling characteristics of streamflow and flow dynamics.

Scaling properties of streamflow was also investigated through spectrum analyses and detrended fluctuation analysis (Özger & Singh, 2013; Telesca et al., 2012). Typically, only drainage area is regarded as an explanatory variable in scaling studies of streamflow. The scaling features of low flows in Northern Iran, for example, were examined as function of drainage area (Modarres, 2010). In recent years, there have been studies on streamflow scaling properties that considered other characteristics besides area. Buttle and Eimers (2009) investigated the scaling behavior of streamflow in terms of watershed characteristics, including area, drainage density, mean slope, and other physiographic properties in a study on south-central Ontario, Canada. They employed simple statistical methods and concluded that the physiography of the basin has affected streamflow depending on the season. Ishak et al. (2011) studied regional flood scaling properties in

New South Wales Australia. They included design rainfall intensities in their study and presented prediction equation for flood quantiles. Farmer et al. (2015) performed a research in southeastern US, working on scaling behavior of daily streamflow considering additional explanatory variables other than drainage area, and evaluate the low flow, and high flow streamflow with more than one explanatory variable. Their results indicated that scaling analyses considering drainage area only brings bias problems to the exponents which was called omitted-variable bias (OVB). To tackle this problem, they suggested using multiple regression considering other catchment characteristics including soil properties, landscape features, slope, and watershed elevation.

Shen and Chui (2021) investigated the scaling characteristic of streamflow and precipitation as a function of temperature across the contiguous United states. They reported that temperature is not an effective predictor of extreme stream flows and its sensitivity. Formetta et al. (2021) investigated peak flow scaling in the United Kingdom using multivariate regression and taking hydro-geomorphological variables and drainage area into account. The results demonstrated that quantile analyses using multiple regression outperformed other methods such as product moments and probability weighted moments. In a global investigation of temporal scaling of streamflow elasticity to precipitation, Zhang et al. (2022) provided insights about streamflow elasticity to precipitation, as well as potential applications of scaling research. First, it demonstrated the relationship between catchment climato-hydrological features and streamflow. Second, scaling laws have the potential to be used as predictive models with some modification and enhancement. They found mean annual precipitation, temperature, aridity index, and evapotranspiration as the major characteristics in the scaling of elasticity. Also, among 5327 catchments around the world, 66% of catchments showed significant scaling of streamflow elasticity to precipitation.

Scaling property in streamflow is important for understanding and managing water resources, since it allows for the prediction of hydrological variables at different scales and the identification of the main drivers of these variables. It also allows for the identification of areas that are at risk of flooding or drought, and to develop strategies for reducing the risk of these events. Additionally, scaling law also facilitates to estimate the potential impacts of climate change on future hydrological conditions (Smith, 1992).

2.3 Streamflow predictions across space and time

Within the field of hydrology, the prediction of streamflow in general refers to the connection between past and future values of streamflow time series. It is one of the most extensively researched topics in water engineering. Several methods have been used to improve prediction, but accurate prediction in space and time remains difficult (Piadeh et al., 2022). Temporal or time scale prediction means forecasting runoff for time periods beyond the measured data. Streamflow prediction across space, or spatial prediction refers to estimating streamflow for catchments where no measurements exist (Sivapalan et al., 2003).

Streamflow has been predicted by using mainly two types of models (Solomatine & Ostfeld, 2008): i) physically-based (process) models which parameterize watershed physical properties (e.g., drainage area, slope, land use/cover, soil properties) and flow routing through drainage networks; ii) empirical and data-driven models providing estimation/prediction of stream flow based on input data of relevant environmental drivers without representing the detailed watershed physical processes (Bourdin et al., 2012).

Statistical models relate streamflow measurements to catchment and climate characteristics. They are developed for gauged catchments and then applied to ungauged areas with similar properties. Regression models provide a simple mathematical transfer function but often have limited physical interpretation. They require quality gauged data to establish robust flow duration curves or other streamflow statistics (Pandey & Nguyen, 1999). The scaling properties of streamflow can be used for predictions across space and time. The distribution of flow events of different magnitude has a characteristic shape, and key scaling parameters can be estimated even with limited data. These parameters relate to dominant runoff mechanisms and may be transferable to other catchments, enabling runoff predictions in ungauged basins or for future periods based on the scaling dynamics of streamflow (Gupta & Dawdy, 1995).

3. MATERIAL AND METHODOLOGY

3.1 Study Area

The Monongahela River Basin is a large river basin located in the north-eastern United States. It covers an area of about 19,000 km^2 and spans across parts of West Virginia, Pennsylvania, and Maryland. The Monongahela River is the main tributary of the Ohio River and is formed by the confluence of the Tygart and West Fork Rivers near Fairmont, West Virginia (Figure 2).

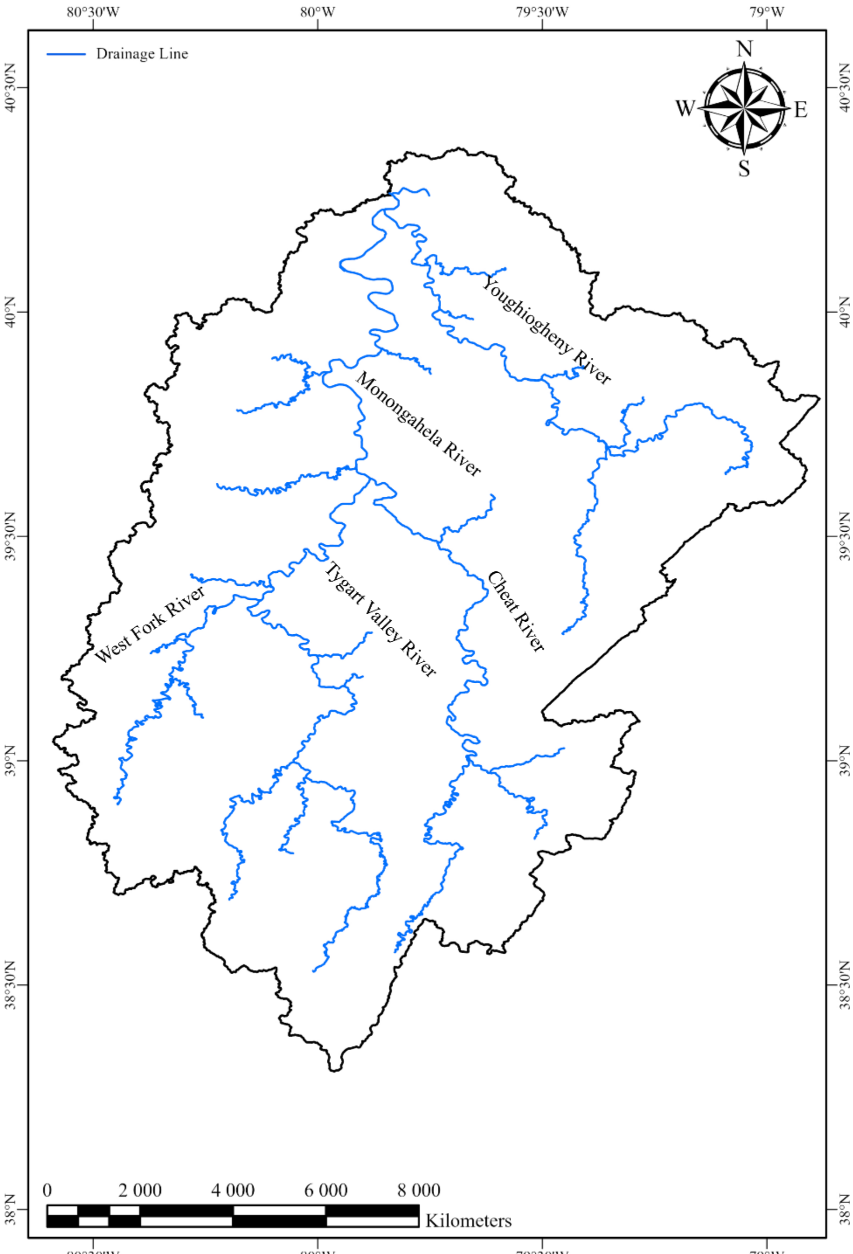

Figure 2. Location of major rivers in Monongahela River Basin

The Monongahela River Basin has a varied range of land uses, including urban, suburban, agricultural, and forested regions (Figure 3). It has a substantial impact on the basin's hydrology, especially has effects on precipitation, evapotranspiration, and runoff. The basins receive an average annual precipitation of about 1000 mm, with the majority of this falling as snow during the winter months. The river's flow is influenced by precipitation, snowmelt, and ground water recharge. The river then flows 206 km north to join the Allegheny and Ohio rivers at the "Point" in Pittsburgh, Pennsylvania (US Army Corps of Engineers, 2011). There are more frequent instances of flooding in the portion of the watershed that is mountainous. Because of the high velocity of the water and the blockage of the floodway caused by debris, the channel capacity is inadequate. Even though

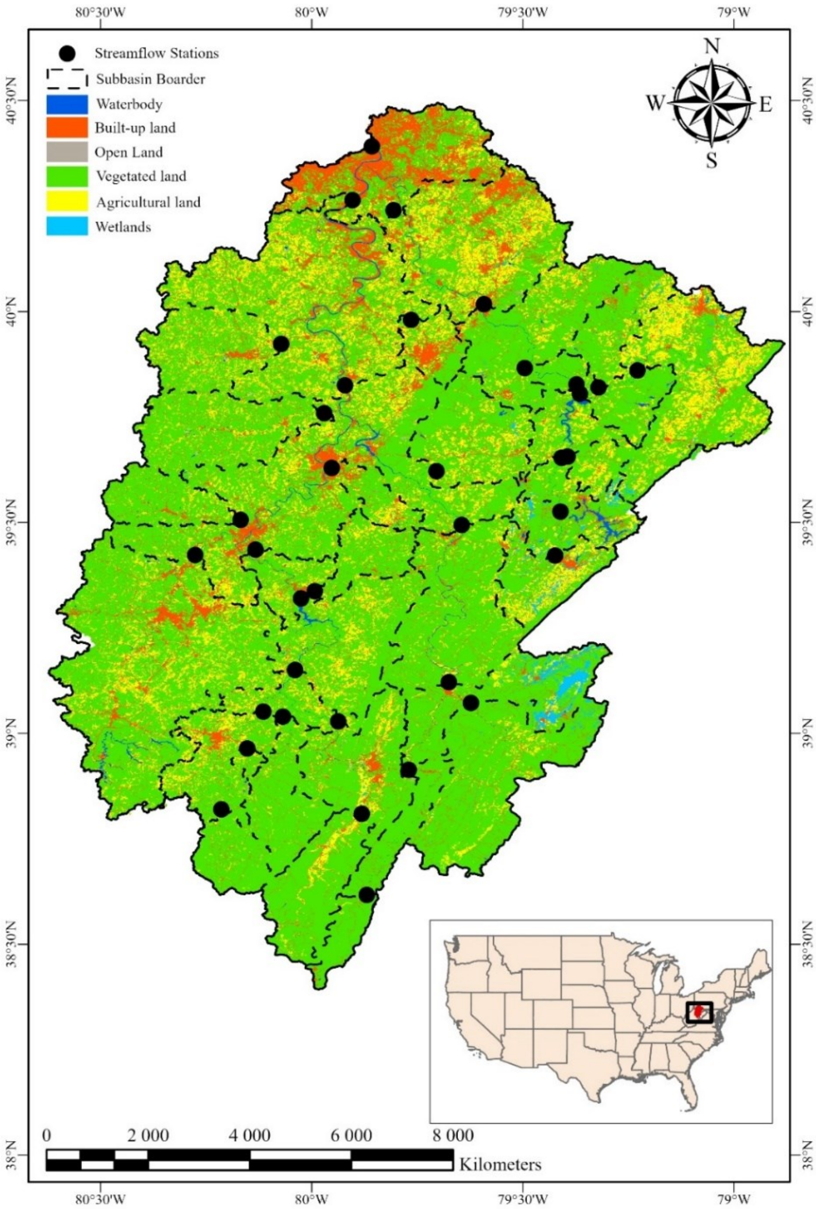

Figure 3. Location of 36 streamflow station across watershed and Landcover

flooding may take place at any time of the year, the months of December through April are often the most hazardous.

The watershed exhibits a humid continental climate, which means because of its location between polar and tropical air masses, it is characterized by unpredictable precipitation patterns and a significant seasonal temperature variation. The average annual precipitation however, is evenly distributed throughout watershed and is approximately 1000 mm. February and august are driest and rainiest months respectively (US Army Corps of Engineers, 2011). The topography of the basin plays an important role in the hydrology of the area. The precipitation and runoff in this region are greatly affected by slope, landscape properties, and elevations due to its mountainous landform.

The Monongahela River Basin's landscape can be characterized as a complex blend of both natural and man-made environments. The natural is characterized by a variety of forest types, including oak-hickory, mixed mesophotic, and northern hardwood forests. The basin also contains a diverse range of wetlands, including swamps, bogs, and fens. The human-altered landscape of the Monongahela River Basin is a result of human activities such as land use changes, urbanization, and industrialization. The basin has a long history of coal mining, which has resulted in the loss of forests and wetlands, as well as acid mine drainage and metal contamination in some of the tributaries. The basin also has a significant amount of water withdrawals for industrial, public, and agricultural uses. Additionally, the Monongahela River has been listed as impaired by the U.S. Environmental Protection Agency (EPA) for various pollutants such as bacteria, mercury, and polychlorinated biphenyls (PCBs). There is limited research and current initiatives and reports about this watershed. There may be a lack of coordinated efforts across agencies, non-profits, local governments, and research institutions to study and address issues in the region. This is because the region is located across three states. Initial assessment for Monongahela River watershed was published by the US Corps of Engineers in September2011 (US Army Corps of Engineers, 2011). Recent updates have been made to the climatological, runoff, and land cover data, which has resulted in more current and accurate information.

3.2 Datasets

The database provided by the United States Geological Survey (USGS) is used to get the streamflow metrics (USGS, 2022). The data provide a time-series ranging from 9 to 30 years and consist of average daily streamflow. The time period covered by the data is from January 1990 to December 2020.Then, pivotal percentiles were computed considering their significance in the flood occurrences. The pivotal percentiles thus provide a reduced but comprehensive representation of the flow duration curve, capturing the full spectrum of hydrologic behaviors within a parsimonious set of 24 values per station. A total of 24 flow percentiles for each streamflow station were selected (Table 2). Every percentile was picked, from 1 to 5, and 95 to 99 percentiles. Then between the 10^{th} and 90^{th} percentiles were selected using ten intervals. In addition, the maximum, minimum, mean streamflow, first, and third quartiles were calculated. Every key flow condition in the station was calculated. The median (50th percentile) represents the mid-point flow, or the flow value that is exceeded 50% of the time. This is an important characteristic of the central tendency of the flow regime. The 25th and 75th percentiles represent the interquartile range, or the range that contains the middle 50% of flows. This indicates the variability of moderate flows. The 5th and 95th percentiles represent the range that contains 90% of flows, indicating the variability of all but the most extreme high and low flows. The 2nd and 98th percentiles provide information on the variability of higher flows that are exceeded only 2% or 2% of the time respectively. Low flow percentiles (2-10%) represent flows in the lower tail of the distribution, indicating baseflow and drought conditions. Common values could indicate regions of similar low flow. High flow percentiles (90-98%) represent flood flows in the upper tail, with common values indicating similar magnitude of floods across regions. Values around 20-30% and 60-70% represent changes in slope of the duration curve, indicating shifts between low-flow, mid-range and high-flow conditions.

The daily (cumulative) observations of historical (1990-2020) precipitation data were obtained from the National Oceanic and Atmospheric Administration (NOAA) of USA (NOAA, 2022). There are 47 climate station across watershed area which record precipitation data. The Thiessen polygon approach, which is a common method for weighing the rain gauge observation according to the streamflow stations, was used in order to assign these climate stations to the streamflow gauges (Figure 4). The more detailed Thiessen polygon approach could be found in (Brassel & Reif, 1979).

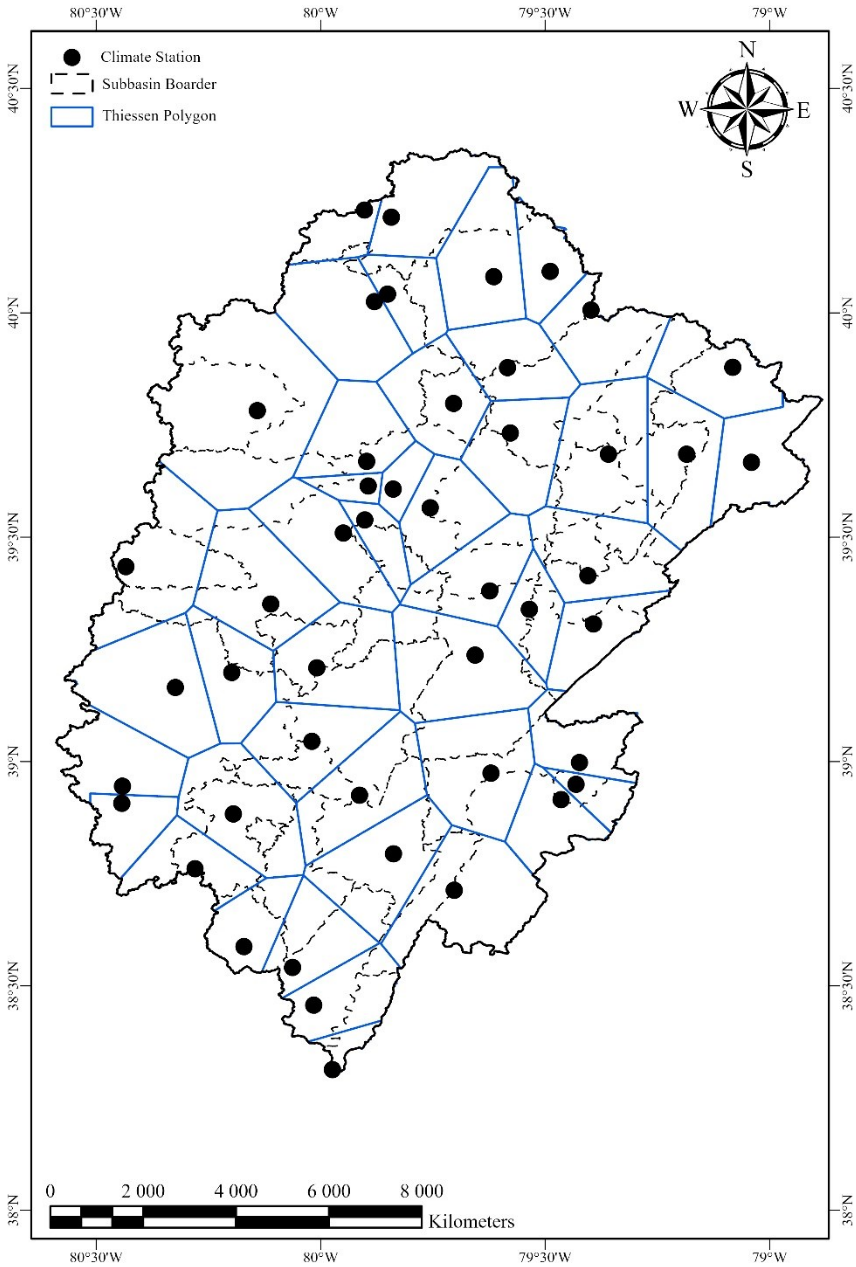

Figure 4. Location of climate stations across watershed and Thiessen polygons

In addition to the maximum and average precipitation, several precipitation percentiles were used to provide a more comprehensive characterization of precipitation beyond just the average and maximum values. Specifically, the 70, 75, 80, 90, 95, and 99 percentiles of precipitation were computed and analyzed. (Table 1).

Table 1. Summary of watershed variables and datasets across the Monongahela River Basin, USA.

Variables	Notations	Units	All-site Mean	Standard deviation	Minimum	Maximum	Coefficient of variation
At-site mean of total daily precipitation	Pmean	mm/day	17.50	11.74	5.17	54.45	0.67
At-site 70th percentile daily precipitation	P70	mm/day	4.42	2.94	1.30	15.22	0.67
At-site 75th percentile daily precipitation	P75	mm/day	6.78	4.56	2.30	25.17	0.67
At-site 80th percentile daily precipitation	P80	mm/day	10.31	7.07	3.80	40.36	0.69
At-site 90th percentile daily precipitation	P90	mm/day	22.95	16.11	9.78	94.01	0.70
At-site 95th percentile daily precipitation	P95	mm/day	36.38	25.67	15.71	149.83	0.71
At-site 99th percentile daily precipitation	P99	mm/day	70.14	49.60	31.35	289.80	0.71
Maximum daily precipitation	Pmax	mm/day	203.67	135.40	82.97	754.07	0.66
Water body	WAT	%	0.25	0.28	0.02	1.31	1.12
Built-up land	BUL	%	4.73	4.68	0.54	24.62	0.99
Open land	OPN	%	0.36	0.41	0.03	2.07	1.12
Vegetated land	VEG	%	48.32	31.53	3.39	94.83	0.65
Agricultural land	AGR	%	8.52	8.02	0.00	30.87	0.94
Wetlands	WET	%	0.60	1.10	0.02	5.79	1.84
Imperviousness	Imp	%	1.38	1.44	0.13	8.49	1.04
Drainage Area	A	Km^2	2330.12	4021.88	36.94	18985.53	1.73
Mean slope	S	%	21.63	6.25	11.37	32.71	0.29
Drainage Density	DD	Km^{-1}	2.20	0.11	1.86	2.46	0.05

Table 2. Summary of streamflow dataset across the Monongahela River Basin, USA.

Variables	Notations	Units	All-site Mean	Standard deviation	Minimum	Maximum	Coefficient of variation
At-site Mean daily stream flow	Q_{mean}	m^3/s	50.89	80.92	0.82	371.30	1.59
At-site Minimum streamflow	Q_{min}	m^3/s	2.93	6.77	0.00	37.94	2.31
At-site 1st percentile streamflow	Q_1	m^3/s	4.95	9.13	0.00	46.16	1.84
At-site 2nd percentile streamflow	Q_2	m^3/s	5.50	9.86	0.00	48.99	1.79
At-site 3rd percentile streamflow	Q_3	m^3/s	5.95	10.46	0.01	51.25	1.76
At-site 4th percentile streamflow	Q_4	m^3/s	6.35	10.99	0.01	53.24	1.73
At-site 5th percentile streamflow	Q_5	m^3/s	6.73	11.51	0.02	55.50	1.71
At-site 10th percentile streamflow	Q_{10}	m^3/s	8.59	14.25	0.03	67.96	1.66
At-site 20th percentile streamflow	Q_{20}	m^3/s	12.16	19.97	0.08	97.58	1.64
At-site 25th percentile streamflow	Q_{25}	m^3/s	14.30	23.52	0.11	115.39	1.65
At-site 30th percentile streamflow	Q_{30}	m^3/s	16.82	28.10	0.15	138.58	1.67
At-site 40th percentile streamflow	Q_{40}	m^3/s	23.14	39.42	0.24	194.54	1.70
At-site 50th percentile streamflow	Q_{50}	m^3/s	31.78	54.78	0.39	269.29	1.72
At-site 60th percentile streamflow	Q_{60}	m^3/s	42.59	72.17	0.57	345.47	1.69
At-site 70th percentile streamflow	Q_{70}	m^3/s	55.52	89.72	0.79	393.60	1.62
At-site 75th percentile streamflow	Q_{75}	m^3/s	65.38	106.20	.93	464.40	1.62
At-site 80th percentile streamflow	Q_{80}	m^3/s	78.12	127.00	1.13	563.51	1.63
At-site 90th percentile streamflow	Q_{90}	m^3/s	118.06	188.41	1.87	841.01	1.60

Table 2. Continue

At-site 95th percentile streamflow	Q_{95}	m^3/s	159.15	244.05	3.00	1087.65	1.53
At-site 96th percentile streamflow	Q_{96}	m^3/s	173.16	261.46	3.45	1169.49	1.51
At-site 97th percentile streamflow	Q_{97}	m^3/s	191.92	285.88	4.05	1277.09	1.49
At-site 98th percentile streamflow	Q_{98}	m^3/s	217.92	320.84	5.07	1438.95	1.47
At-site 99th percentile streamflow	Q_{99}	m^3/s	271.80	389.19	6.99	1730.44	1.43
At-site Maximum streamflow	Q_{max}	m^3/s	803.62	1159.32	37.38	5323.57	1.44

Comparing multiple percentiles provides more nuanced information than any single value. For example, increasing 70 but decreasing 95 percentiles could indicate a shift to more frequent lighter rains and fewer intense storms. The 70-80 percentiles represent moderate precipitation amounts that are exceeded 70-80% of the time. Analyzing these together provides insight into typical rainfall intensities and the variability of average storms. The 80-99 percentiles: Represent heavier precipitation that is exceeded only 5-15% of the time. Analyzing these higher percentiles indicates the intensity of intense storms and extremes.

In order to delineate watershed and export its characteristics. ArcGIS Pro 2.5 and 1/3 arc-second (approximately 10 m) resolution digital elevation models were used. This includes the mean slope, imperviousness, drainage density, and area. Drainage density is calculated by dividing the entire length of a channel within a drainage basin by its total area. The imperviousness data is obtained from the Multi-Resolution Land Characteristics (MRLC) consortium (MRLC, 2022). The data is provided at regular intervals, once every two years between the years 2001 and 2019. After collecting the data, the mean value was computed for use in regression models. National Land Cover Database (NLCD) categorizes land covers into six broad groups: (1) waterbody (WAT), (2) built-up land (BUL), (3) open land (OPN), (4) vegetated land (VEG), (5) agricultural land (AGR, and (6) wetlands (WTL). Land cover data are also provided by MRLC for the same time period as imperviousness data.

A nested approach was used to work with watershed data including climatic, and hydrologic data (e.g., Area, precipitation, slope, imperviousness, land cove, drainage density) (Text S1, Appendix). There are some key benefits to using a nested or cumulative approach. First it captures scaling at multiple spatial scales. It is possible to identify how scaling patterns change as the size of the watershed increases. This provides insights into how processes influencing streamflow may dominate at different scales. Also, hydrological processes that shape streamflow can vary in importance across scales. A nested approach ensures you capture the primary processes driving scaling at each individual scale, not just an overall scaling pattern averaged across scales (Ayalew et al., 2015). However nested approach which will inherently involve spatial autocorrelations (i.e., pseudoreplication)(Kroll et al., 2014). A Monte-Carlo bootstrap approach was used to randomly resample the data 1,000 times (while maintaining the same number of stations, n=36 in each iteration) in order to break the underlying spatial correlations in the data set and eliminate the influence of pseudoreplication model estimations. The number of iterations used (1000) in this Monte Carlo bootstrap approach is considered an adequately large and robust number that helps ensure the results and conclusions drawn from them are stable, statistically sound, unbiased, and computable within reasonable time and resource constraints (Abdul-Aziz & Gebreslase, 2023).

There are three approaches for doing a scaling assessment. Quantile analyses, probability weighted moments, and product moments (Formetta et al., 2021). Quantile analyses is a statistical method used to analyze the distribution of a variable, such as streamflow, by dividing the data into quantiles, or intervals that contain a certain percentage of the data. The intervals are determined by the quantile values, which are calculated by dividing the data into equal parts. Quantile analyses method was first introduced by Benson (1962) and has been widely used in literature for regional flood frequency analyses. By analyzing the quantiles of the data, it is possible to identify patterns in the data that may not be immediately obvious, such as nonlinear relationships or outliers (Gupta et al., 1994). In this method, each flood quantile can be expressed as a regression function of explanatory variables which are watershed characteristics. It can be written as:

$$Q_T = f(A, P, S, \ldots) \quad (1)$$

Where Q_T is flow at quantile T, and f is function of basin characteristics (such as drainage area (A), precipitation (P), channel slope (S)). It's important to note that quantile analysis is a descriptive method, and it doesn't provide a causal explanation of the relationship between variables. Also, it's important to check the assumptions of normality and homoscedasticity before applying quantile analysis, as it relies on these assumptions (Pandey & Nguyen, 1999).

3.3 Data Analytics to establish the relative streamflow controls

After gathering data from different sources, the data must be prepared thoroughly for unbiased analysis. Three steps were followed before using the data in power law scaling analyses. The steps ensured there would be no bias in the modeling and interpretation. By applying the results of Pearson correlation matrix, patterns in data that are immediately visible can be found, such as relationships between different hydrological variables that are directly connected to streamflow. However, there was still possibility of multicollinearity (mutual correlations among the drivers) in our model (Abdul-Aziz et al., 2018; Abdul-Aziz & Ahmed, 2017; Al-Rawas & Valeo, 2010). This issue had the potential to impact analyses in different ways. First, it is difficult to determine which variables have the most significant impact on the streamflow. The second problem is that the output model does not have a robust scaling exponent, and the third problem is that there is a possibility of overfitting, which results in misleading performance.

In order to resolve the issue of multicollinearity principal component analysis (PCA) was used (Jolliffe, 2002). This technique first standardize data by subtracting the mean and dividing by the standard deviation, then calculating the covariance matrix of the data. The covariance matrix was then used to construct eigenvectors and eigenvalues, and the eigenvectors with the highest eigenvalues are chosen as the principal components (PCs). The principal components refer to the linear combinations of the original variables that explain the majority of the variation in the data. Therefore, the total number of PCs was equal to the initial variable number. Such analyses can be used as a data compression technique by keeping only the first few principal components that capture most of the variance in the data and discarding the rest. This can significantly reduce the size of the data without sacrificing much information.

3.4 Estimation of Emergent Power Law Scaling Model

In this section the relationship between streamflow quantiles and watershed characteristics as a power law scaling function as summarized. A spatial scaling across Monongahela River basin was performed.

Power law scaling have been widely used in hydrological analyses and predictions(Vogel et al., 1999). To predict streamflow Q from n distinct streamflow predictors $\mathrm{X_i}(i = 1, \ldots, n)$, a multiple power law equation. Eq. (2) was employed then is logarithmically transformed to a linear model with coefficients $a_\mathrm{i}(i = 0, \ldots, n)$. It can be generated by multiple linear regression. This model is shown by the Eq (2).

$$Q = 10^{a_0}.{X_1}^{a_1}.{X_2}^{a_2}.{X_3}^{a_3} \ldots {X_n}^{a_n} \tag{2}$$

$$Log(Q) = a_0 + a_1.Log(X_1) + a_2.Log(X_2) + a_3.Log(X_3) + \cdots + a_n.Log(X_n) \tag{3}$$

The multiple linear regression was performed by ordinary least squares (OLS) technique. The p-values of the independent variables were calculated to assess the significance of each predictor in the regression analysis. The coefficients' *p* values show whether or not these associations are statistically significant. As the p-value decreases, the evidence against the null hypothesis becomes stronger.

Power law scaling was applied to streamflow data from 36 gauging stations across the watershed, using the flows as the dependent variables. In addition, 11 distinct watershed features were selected as dependent variables for regression analysis. Adopting a cumulative approach for each sub-watershed, regression was performed using streamflow data from individual sub-basins. The regression technique was applied to the streamflow data within each sub-watershed. This strategy is effective because it ensures that all of the sub-watersheds are linked to one another and are subject to the influence of the upstream sub-watershed. The data was log-transformed to provide a normalized scale, facilitating analysis. Once the data was in log10 form, tests for autocorrelation revealed significant correlations between several key variables. By converting the raw data to logs, the normalized scale allowed for detection of correlations that might otherwise be obscured. The log-transformed data provided insights that would not have been possible using the raw values. Even if the parameter estimators will be unbiased, the failure of our linear regression model to meet the autocorrelation criterion may have spurious results (Mcguirk & Spanos, 2002). Autocorrelation, which is common in time series data, can result in a strong correlation between variables that are erroneously related. Therefore, a bootstrap resampling approach was used with 1000 iterations to select data panels at random (keep the sample size in each set) for model estimations to break the spatial correlations of each variable. For more accurate model estimate, a cross-validation mechanism was implemented. In this technique, the data is simply split into several sets for modeling and prediction. To determines the percentage of the data that goes into each subset, there are some things to consider include: statistical power and using a larger percentage of the data for training (e.g., 70%) that will result in models with higher statistical power to detect patterns. Also, the testing sets still need enough data (e.g. 30%) to provide a reliable estimate of model performance, or power will be low (Kuhn & Johnson, 2013b). In each iteration, 70 % of the data was used for model calibration (training) and 30% for validation (test). This approach was used to simulate ungauged basins in the analyses, enabling evaluation of the regression models' effectiveness on unmonitored streams. All the analyses and modeling were performed with Python and Microsoft Excel.

To choose the important predictor in power law equation a forward selection approach was used (Kuhn & Johnson, 2013a). The main idea behind forward selection is to start with an empty model and add one predictor variable at a time based on statistical criterion (in our case, PCA, *p* value, model parsimony and model efficiency) until no further improvement in the fit of the model can be achieved. Then the final model was evaluated and checked for the validity of the assumptions of multiple linear regression, such as linearity, independence, homoscedasticity, and normality of residuals.

The Akaike Information Criterion (AIC) (Parzen & Priestley, 1984) is a statistical criterion for model selection that allows for the comparison of the relative quality of different models. The main idea behind AIC is to balance the goodness of fit of a model with its complexity.

$$AIC(p) = \ln\left(\frac{SSE}{n}\right) + \frac{2p}{n} \tag{4}$$

where p = number of components; n = total number of observations; SSE = total sum of squared errors = $\sum_{i=1}^{n}(Q_{i,o} - Q_{i,p})^2$. $Q_{i,o}$ is observed discharge and $Q_{i,p}$ predicted discharge.

The model efficiency was measured by two indicators, Nash-Sutcliffe efficiency (NSE) and root-mean-square error to the standard deviation of observations (RSR).

NSE was computed as follows:

$$NSE = 1 - \frac{\sum_{i=1}^{T}(Q_{i,o} - Q_{i,p})^2}{\sum_{i=1}^{T}(Q_{i,o} - \overline{Q_o})^2} \tag{5}$$

Where $\overline{Q_o}$is the mean of observed discharges, $Q_p$is predicted discharge, and $Q_{i,o}$is observed discharge at time t. When a model is accurate, and the estimation error variance is zero, the resulting Nash–Sutcliffe Efficiency is 1 (NSE = 1). Alternatively, NSE = 0 when the estimation error variance is equal to the variance of the observed time series. Negative NSE means than the mean observation performs better than model (Nash & Sutcliffe, 1970).

Also, the ratio of the root-mean-square error to the standard deviation of observations (RSR) was calculated as:

$$RSR = \frac{\sqrt{\sum_{i=1}^{N}\frac{\left(Y_{i,o} - Y_{i,p}\right)^2}{N}}}{\sigma_o} \tag{6}$$

where N is the total number of observations; $Y_{i,p}$ and $Y_{i,o}$are the i-th model prediction and the corresponding observation, respectively; and σ_o is the standard deviation of observations. RSR contains the advantages of error index statistics and adds a normalization factor using the observations standard deviation, allowing the reported values to be applicable to varied components. The suggested RSR value of less than 0.5 is regarded as the strictest ("very good") grade, whereas "good" and "satisfactory" correspond to scores of 0.6 and 0.7, respectively (Moriasi et al., 2007).

4. RESULTS

4.1 Correlations of Flow Regimes and Environmental Variables

The interconnections between streamflow percentiles with meteorological drivers, and hydrologic components were quantified via Pearson correlation (Table 3, 4, and 5). The maximum and minimum streamflow percentiles had the lowest correlation between them (r = 0.80, *p* value = 0.00), though it was still highly correlated. Streamflow percentiles and precipitation percentiles were moderately to strongly correlated. The maximum precipitation and streamflow at 70% had the strongest correlation (r = 0.7, *p* value = 0.00). Additionally, the mean precipitation and minimum streamflow had the lowest correlations among precipitation and streamflow percentiles (r = 0.14, *p* value = 0.00). There was a very strong correlation between all of the precipitation percentiles with each other. Also, there was a very strong correlation between watershed area and streamflow percentiles (r = 0.85-0.99, *p* value = 0.00- 0.05). There was a fairly significant correlation between land cover features and drainage density, mean slope, or imperviousness among the other watershed parameters. It demonstrated multicollinearity between watershed climato-hydrological parameters and land cover characteristics. For instance, there was a significant link between mean slope and wetland land cover (r = -0.69, *p* value = 0.00). Furthermore, there were weak correlations between mean slope, drainage density, imperviousness, and streamflow percentiles (r=0.02 to 0.16, *p* value= 0.00 to 0.78).

Table 3. Pearson Correlation Coefficients of Streamflow regimes with the hydrological, and land cover variables

Variables	Q_{mean}	Q_{min}	Q_1	Q_2	Q_3	Q_4	Q_5	Q_{10}	Q_{20}	Q_{30}	Q_{40}	Q_{50}	Q_{60}	Q_{70}	Q_{80}	Q_{90}	Q_{95}	Q_{96}	Q_{97}	Q_{98}	Q_{99}	Q_{max}
WAT	-0.24	-0.18	-0.19	-0.19	-0.19	-0.19	-0.19	-0.20	-0.21	-0.22	-0.22	-0.23	-0.24	-0.24	-0.24	-0.25	-0.25	-0.25	-0.25	-0.25	-0.26	-0.25
BUL	-0.41	-0.32	-0.38	-0.39	-0.39	-0.40	-0.40	-0.41	-0.41	-0.40	-0.39	-0.39	-0.39	-0.41	-0.41	-0.41	-0.42	-0.42	-0.42	-0.42	-0.42	-0.39
OPN	-0.28	-0.24	-0.28	-0.29	-0.29	-0.29	-0.29	-0.29	-0.29	-0.28	-0.27	-0.28	-0.28	-0.29	-0.29	-0.29	-0.28	-0.28	-0.28	-0.28	-0.27	-0.19
VEG	-0.63	-0.53	-0.63	-0.64	-0.65	-0.65	-0.66	-0.67	-0.65	-0.63	-0.61	-0.60	-0.61	-0.63	-0.62	-0.63	-0.63	-0.63	-0.63	-0.63	-0.63	-0.56
AGR	-0.41	-0.31	-0.36	-0.37	-0.38	-0.38	-0.39	-0.40	-0.41	-0.40	-0.40	-0.39	-0.40	-0.41	-0.41	-0.41	-0.42	-0.42	-0.42	-0.42	-0.43	-0.38
WTL	-0.18	-0.15	-0.18	-0.18	-0.18	-0.18	-0.18	-0.18	-0.17	-0.17	-0.17	-0.18	-0.18	-0.19	-0.19	-0.19	-0.19	-0.19	-0.19	-0.19	-0.19	-0.15
Imp	0.10	0.14	0.13	0.13	0.12	0.12	0.12	0.11	0.11	0.11	0.11	0.11	0.10	0.09	0.10	0.09	0.09	0.09	0.08	0.08	0.08	0.08
A	1.00	0.85	0.92	0.93	0.94	0.94	0.94	0.96	0.98	0.99	0.99	1.00	1.00	0.99	1.00	1.00	0.99	0.99	0.99	0.99	0.99	0.97
S	0.09	-0.07	-0.08	-0.08	-0.08	-0.07	-0.07	-0.05	0.00	0.04	0.06	0.08	0.09	0.10	0.10	0.10	0.11	0.11	0.12	0.12	0.13	0.11
DD	0.15	0.10	0.12	0.12	0.12	0.12	0.12	0.13	0.13	0.13	0.13	0.14	0.14	0.15	0.15	0.15	0.16	0.16	0.16	0.16	0.16	0.13

Note. Q_{mean}, Q_{min}, Q_{max}, WAT, BUL, OPN, VEG, AGR, WTL, Imp, A, S, DD respectively refer to at-site mean, minimum, and maximum streamflow, water bodies, built up lands, open lands, vegetated lands, agricultural lands, wetlands, imperviousness, area, mean slope, and drainage density. The rest of the variables represent percentiles of streamflow

Table 4. Pearson Correlation Coefficients of Streamflow regimes with the precipitation percentiles

Variables	Q_{mean}	Q_{min}	Q_1	Q_2	Q_3	Q_4	Q_5	Q_{10}	Q_{20}	Q_{30}	Q_{40}	Q_{50}	Q_{60}	Q_{70}	Q_{80}	Q_{90}	Q_{95}	Q_{96}	Q_{97}	Q_{98}	Q_{99}	Q_{max}
P_{mean}	0.35	0.14	0.27	0.29	0.30	0.31	0.31	0.33	0.33	0.33	0.32	0.31	0.32	0.35	0.35	0.35	0.36	0.36	0.36	0.36	0.36	0.31
P_{70}	0.43	0.17	0.31	0.33	0.34	0.35	0.36	0.37	0.39	0.39	0.39	0.39	0.40	0.44	0.44	0.44	0.45	0.45	0.45	0.44	0.45	0.38
P_{75}	0.50	0.21	0.36	0.38	0.40	0.41	0.41	0.43	0.45	0.46	0.46	0.45	0.47	0.51	0.51	0.51	0.51	0.52	0.51	0.51	0.52	0.44
P_{80}	0.55	0.25	0.41	0.42	0.44	0.45	0.46	0.48	0.50	0.51	0.51	0.51	0.53	0.57	0.57	0.57	0.57	0.57	0.57	0.57	0.57	0.50
P_{90}	0.61	0.31	0.47	0.49	0.50	0.51	0.52	0.54	0.56	0.57	0.57	0.57	0.59	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.55
P_{95}	0.63	0.33	0.49	0.51	0.52	0.53	0.54	0.56	0.58	0.59	0.59	0.59	0.61	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.57
P_{99}	0.65	0.35	0.51	0.53	0.54	0.55	0.56	0.58	0.60	0.60	0.61	0.60	0.62	0.67	0.66	0.66	0.67	0.67	0.67	0.66	0.66	0.58
P_{max}	0.69	0.41	0.57	0.58	0.60	0.61	0.61	0.63	0.64	0.65	0.65	0.65	0.66	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.61

Note. Q_{mean}, Q_{min}, Q_{max}, P_{mean} , P_{min} , P_{max} respectively refer to at-site mean, minimum, and maximum streamflow, and at-site mean, minimum, and maximum daily precipitation. The rest of the variables represent percentiles of streamflow and daily precipitation.

Table 5. Mutual Pearson Correlation Coefficients of the hydrological, and land cover variables

Variables	WAT	BUL	OPN	VEG	AGR	WTL	Imp	A	S	DD
WAT	1.00	0.07	0.00	-0.09	0.25	0.65	-0.01	-0.23	-0.60	0.03
BUL	0.07	1.00	0.24	0.42	0.70	0.12	0.78	-0.38	-0.27	0.16
OPN	0.00	0.24	1.00	0.34	0.47	0.20	-0.02	-0.28	-0.18	-0.04
VEG	-0.09	0.42	0.34	1.00	0.38	-0.03	-0.15	-0.59	0.25	-0.32
AGR	0.25	0.70	0.47	0.38	1.00	0.38	0.35	-0.38	-0.47	0.14
WTL	0.65	0.12	0.20	-0.03	0.38	1.00	-0.01	-0.19	-0.51	-0.02
Imp	-0.01	0.78	-0.02	-0.15	0.35	-0.01	1.00	0.12	-0.33	0.33
A	-0.23	-0.38	-0.28	-0.59	-0.38	-0.19	0.12	1.00	0.09	0.15
S	-0.60	-0. 27	-0.18	0.25	-0.47	-0.51	-0.33	0.09	1.00	0.08
DD	0.03	0.16	-0.04	-0.32	0.14	-0.02	0.33	0.15	0.08	1.00

Note. WAT, BUL, OPN, VEG, AGR, WTL, Imp, A, S, DD respectively refer to water bodies, built up lands, open lands, vegetated lands, agricultural lands, wetlands, imperviousness, area, mean slope, and drainage density.

The development of an accurate power-law model requires an understanding of the mutual correlation that exists between the variables.

4.2 Relative Orientations of Flow Regimes and Environmental Variables

The relative orientation of environmental drivers and flow regimes was investigated with three PCA (Table 6, 7, 8, and 9). These analyses revealed collinearity correlation information as well as information on significant variables in a scaling model. Three sets of analyses were performed in which in the first one, all the hydrological and land cover variables were associated. In the second analysis, climatic variables including precipitation percentiles were investigated. In the last analysis both climatic and hydrological variables were incorporated in PCA. Based on the results, the area showed a strong linkage with all streamflow percentiles in all analysis. Also, AGU, BUL, WTL, and WAT formed a land cover component with strong positive linkage together. Additionally, vegetation land cover showed a strong negative linkage with flow regimes and slope, and imperviousness orientated rather orthogonally with flow regimes, indicating their weak control.

Table 6. Mean of loadings and variance (in parenthesis) of all hydrological, and land cover variables on each principal component (PC) across all study sites.

Variables	PC1 (33%)	PC2 (21%)	PC3 (16%)	PC4 (9%)	PC5 (8%)	PC6 (5%)	PC7 (3%)	PC8 (3%)	PC9 (1%)	PC10 (0%)	PC11 (0%)
Q_{mean}	0.42	0.29	0.05	0.30	0.06	0.32	-0.08	-0.03	-0.07	-0.17	0.71
WAT	-0.22	0.24	-0.44	-0.32	-0.05	0.31	-0.64	0.26	0.13	0.00	0.02
BUL	-0.38	0.19	0.42	-0.08	0.19	0.16	0.07	0.15	0.01	-0.73	-0.10
OPN	-0.27	-0.07	0.01	0.75	-0.31	-0.22	-0.21	0.42	0.07	-0.01	-0.01
VEG	-0.32	-0.41	0.16	0.07	0.09	0.47	-0.12	-0.04	-0.63	0.24	0.08
AGR	-0.42	0.18	0.12	0.22	-0.09	0.18	-0.07	-0.69	0.41	0.19	0.04
WTL	-0.23	0.24	-0.46	0.08	-0.15	0.34	0.70	0.18	-0.04	0.05	-0.01
Imp	-0.13	0.44	0.45	-0.11	0.28	-0.07	0.07	0.38	0.06	0.58	0.08
A	0.41	0.29	0.07	0.31	0.07	0.37	-0.12	-0.06	-0.11	0.04	-0.69
S	0.20	-0.43	0.30	-0.13	-0.30	0.46	0.10	0.27	0.53	0.07	0.00
DD	0.03	0.30	0.27	-0.25	-0.81	-0.07	-0.03	-0.08	-0.33	-0.01	0.00

Note. Q_{mean}, WAT, BUL, OPN, VEG, AGR, WTL, Imp, A, S, DD respectively refer to at-site mean streamflow, water bodies, built up lands, open lands, vegetated lands, agricultural lands, wetlands, imperviousness, area, mean slope, and drainage density.

Table 7. Mean of loadings and variance (in parenthesis) of all hydrological, and land cover variables on each principal component (PC) across all study sites.

Variables	PC1 (33%)	PC2 (21%)	PC3 (16%)	PC4 (9%)	PC5 (8%)	PC6 (5%)	PC7 (3%)	PC8 (3%)	PC9 (1%)	PC10 (0%)	PC11 (0%)
Q_{50}	0.41	0.30	0.05	0.32	0.06	0.32	-0.08	-0.03	-0.11	-0.13	-0.70
WAT	-0.22	0.24	-0.44	-0.31	-0.03	0.33	-0.63	0.26	0.14	0.00	-0.02
BUL	-0.38	0.19	0.42	-0.06	0.20	0.16	0.07	0.15	0.01	-0.74	0.06
OPN	-0.27	-0.07	0.01	0.72	-0.36	-0.23	-0.21	0.42	0.07	-0.01	0.01
VEG	-0.32	-0.41	0.16	0.10	0.10	0.47	-0.11	-0.04	-0.62	0.25	-0.02
AGR	-0.42	0.18	0.12	0.22	-0.10	0.18	-0.06	-0.68	0.42	0.19	-0.05
WTL	-0.24	0.24	-0.46	0.07	-0.15	0.33	0.71	0.18	-0.04	0.05	0.01
Imp	-0.13	0.45	0.45	-0.09	0.28	-0.08	0.07	0.38	0.06	0.58	-0.05
A	0.41	0.30	0.07	0.31	0.05	0.35	-0.11	-0.05	-0.06	0.00	0.70
S	0.20	-0.43	0.30	-0.14	-0.28	0.47	0.10	0.27	0.53	0.06	-0.03
DD	0.03	0.30	0.27	-0.30	-0.79	-0.04	-0.03	-0.09	-0.33	-0.01	0.00

Note. Q_{50} , WAT, BUL, OPN, VEG, AGR, WTL, Imp, A, S, DD respectively refer to at-site median streamflow, water bodies, built up lands, open lands, vegetated lands, agricultural lands, wetlands, imperviousness, area, mean slope, and drainage density.

Table 8. Mean of loadings and variance (in parenthesis) of all hydrological, and land cover variables on each principal component (PC) across all study sites.

Variables	PC1 (33%)	PC2 (21%)	PC3 (16%)	PC4 (9%)	PC5 (8%)	PC6 (5%)	PC7 (3%)	PC8 (3%)	PC9 (1%)	PC10 (0%)	PC11 (0%)
Q_{min}	0.36	0.34	0.05	0.38	0.03	0.35	-0.16	-0.10	-0.51	-0.44	-0.03
WAT	-0.23	0.24	-0.44	-0.32	0.07	0.31	-0.58	0.36	0.12	-0.09	0.01
BUL	-0.40	0.16	0.42	-0.03	0.22	0.12	0.09	0.14	-0.04	-0.03	-0.74
OPN	-0.28	-0.08	0.00	0.61	-0.53	-0.21	-0.14	0.45	0.02	-0.06	-0.02
VEG	-0.32	-0.41	0.15	0.13	0.12	0.48	-0.13	-0.03	-0.40	0.44	0.26
AGR	-0.44	0.17	0.12	0.18	-0.13	0.20	-0.11	-0.62	0.42	-0.24	0.19
WTL	-0.25	0.23	-0.46	-0.01	-0.15	0.30	0.73	0.09	-0.11	-0.01	0.04
Imp	-0.14	0.43	0.46	-0.04	0.27	-0.11	0.13	0.37	0.02	-0.04	0.58
A	0.40	0.30	0.08	0.28	-0.04	0.37	0.00	0.09	0.45	0.55	-0.10
S	0.20	-0.44	0.29	-0.21	-0.21	0.45	0.16	0.28	0.29	-0.44	0.06
DD	0.02	0.28	0.28	-0.46	-0.70	0.02	-0.09	-0.12	-0.28	0.20	-0.01

Note. Q_{min} , WAT, BUL, OPN, VEG, AGR, WTL, Imp, A, S, DD respectively refer to at-site minimum streamflow, water bodies, built up lands, open lands, vegetated lands, agricultural lands, wetlands, imperviousness, area, mean slope, and drainage density.

Table 9. Mean of loadings and variance (in parenthesis) of all hydrological, and land cover variables on each principal component (PC) across all study sites.

Variables	PC1 (33%)	PC2 (21%)	PC3 (16%)	PC4 (9%)	PC5 (8%)	PC6 (5%)	PC7 (3%)	PC8 (3%)	PC9 (1%)	PC10 (0%)	PC11 (0%)
Q_{max}	0.40	0.30	0.05	0.37	0.05	0.32	-0.08	-0.01	-0.08	0.70	-0.06
WAT	-0.23	0.24	-0.45	-0.28	-0.06	0.35	-0.65	0.23	0.13	0.01	0.01
BUL	-0.39	0.17	0.42	-0.06	0.19	0.17	0.06	0.15	0.01	0.01	-0.74
OPN	-0.26	-0.06	0.01	0.73	-0.30	-0.27	-0.22	0.41	0.08	-0.08	-0.01
VEG	-0.31	-0.42	0.16	0.12	0.08	0.46	-0.11	-0.05	-0.62	0.00	0.25
AGR	-0.43	0.17	0.12	0.22	-0.09	0.15	-0.04	-0.70	0.41	0.04	0.19
WTL	-0.24	0.24	-0.46	0.10	-0.15	0.33	0.70	0.19	-0.04	-0.04	0.05
Imp	-0.14	0.44	0.45	-0.12	0.28	-0.05	0.05	0.38	0.06	-0.01	0.58
A	0.41	0.31	0.07	0.29	0.08	0.33	-0.11	-0.11	-0.09	-0.71	-0.07
S	0.21	-0.42	0.30	-0.10	-0.30	0.47	0.09	0.26	0.54	-0.03	0.07
DD	0.03	0.30	0.27	-0.25	-0.81	-0.06	-0.03	-0.08	-0.33	0.01	-0.01

Note. Q_{max}, WAT, BUL, OPN, VEG, AGR, WTL, Imp, A, S, DD respectively refer to at-site maximum streamflow, water bodies, built up lands, open lands, vegetated lands, agricultural lands, wetlands, imperviousness, area, mean slope, and drainage density.

In the first analyses all the hydrological and land cover variables loading were represented. Based on the results, first three PCs which are linear combination of all environmental variables explained, approximately 33%,21%, and 16% of data variance respectively on average for all the flow regimes. According to the first two PCs orientation (Figure S1 in Appendix), area had the strongest linkage with flow regimes (loading score = 0.41). Furthermore, the low drainage density loading (0.03) on PC1 and PC2 suggested a weak relation with flow regimes.

In the second analysis, PC loadings of climatological variables and flow regimes was calculated (Table 10-13). Approximately 88% and 10% of the total data variance were explained, respectively, by the first two PCs (PC1, PC2) (Figure S2 in Appendix). There was a moderate linkage between flow regimes and precipitation percentiles based on loading rates (0.32-0.35). Furthermore, the relative orientation of different precipitation percentiles suggested a nearly identical connection between these variables. Among different percentiles, maximum, and mean precipitation suggested potential positive correlation with mean streamflow (loading rate 0.35, and 0.32 respectively).

Table 10. Mean of loadings and variance (in parenthesis) of all climatological variables on each principal component (PC) across all study sites.

Variables	PC1 (88%)	PC2 (10%)	PC3 (2%)	PC4 (1%)	PC5 (0%)	PC6 (0%)	PC7 (0%)	PC8 (0%)	PC9 (0%)
Q_{mean}	0.22	0.85	0.44	-0.19	0.03	0.01	0.00	0.00	0.00
P_{mean}	0.32	-0.37	0.67	0.19	-0.47	0.19	-0.14	0.05	0.00
P_{70}	0.34	-0.27	0.24	-0.23	0.44	-0.60	-0.03	-0.14	-0.36
P_{75}	0.35	-0.17	0.04	-0.21	0.27	0.20	0.53	-0.13	0.63
P_{80}	0.35	-0.07	-0.17	-0.23	0.31	0.38	-0.49	0.55	0.00
P_{90}	0.35	0.02	-0.27	-0.13	-0.18	0.40	0.40	-0.17	-0.63
P_{95}	0.35	0.06	-0.32	-0.08	-0.23	-0.06	-0.51	-0.63	0.22
P_{99}	0.35	0.09	-0.30	-0.03	-0.49	-0.51	0.19	0.47	0.15
P_{max}	0.34	0.15	-0.10	0.87	0.31	-0.02	0.04	0.00	-0.01

Note. Q_{mean}, P_{mean}, P_{max} respectively refer to at-site mean streamflow, at-site mean, and maximum daily precipitation. The rest of the variables represent percentiles of at-site daily precipitation.

Table 11. Mean of loadings and variance (in parenthesis) of all climatological variables on each principal component (PC) across all study sites.

Variables	PC1 (88%)	PC2 (9%)	PC3 (2%)	PC4 (1%)	PC5 (0%)	PC6 (0%)	PC7 (0%)	PC8 (0%)	PC9 (0%)
Q_{min}	0.11	0.94	0.30	-0.14	0.03	0.01	0.00	0.00	0.00
P_{mean}	0.33	-0.21	0.70	0.25	-0.48	0.19	-0.13	0.05	0.00
P_{70}	0.35	-0.18	0.32	-0.21	0.44	-0.59	-0.06	-0.17	-0.35
P_{75}	0.36	-0.12	0.11	-0.21	0.27	0.18	0.53	-0.09	0.64
P_{80}	0.36	-0.06	-0.11	-0.25	0.31	0.39	-0.48	0.56	-0.03
P_{90}	0.36	0.01	-0.25	-0.15	-0.18	0.37	0.45	-0.16	-0.63
P_{95}	0.36	0.04	-0.30	-0.13	-0.21	0.01	-0.50	-0.64	0.23
P_{99}	0.35	0.06	-0.31	-0.07	-0.48	-0.54	0.14	0.46	0.15
P_{max}	0.35	0.14	-0.20	0.85	0.31	-0.02	0.03	0.00	-0.01

Note. Q_{50}, P_{mean}, P_{max} respectively refer to at-site median streamflow, at-site mean, and maximum daily precipitation. The rest of the variables represent percentiles of at-site daily precipitation.

Table 12. Mean of loadings and variance (in parenthesis) of all climatological variables on each principal component (PC) across all study sites.

Variables	PC1 (86%)	PC2 (11%)	PC3 (2%)	PC4 (1%)	PC5 (0%)	PC6 (0%)	PC7 (0%)	PC8 (0%)	PC9 (0%)
Q_{50}	0.20	0.87	0.42	-0.18	0.03	0.01	0.00	0.00	0.00
P_{mean}	0.32	-0.35	0.67	0.21	-0.48	0.19	-0.14	0.05	0.00
P_{70}	0.34	-0.26	0.26	-0.23	0.43	-0.60	-0.03	-0.14	-0.36
P_{75}	0.35	-0.16	0.06	-0.21	0.27	0.20	0.52	-0.13	0.63
P_{80}	0.35	-0.07	-0.16	-0.24	0.31	0.38	-0.49	0.56	-0.01
P_{90}	0.35	0.02	-0.26	-0.13	-0.18	0.40	0.40	-0.17	-0.63
P_{95}	0.35	0.06	-0.32	-0.09	-0.22	-0.05	-0.51	-0.63	0.22
P_{99}	0.35	0.08	-0.31	-0.04	-0.49	-0.51	0.19	0.47	0.15
P_{max}	0.35	0.15	-0.12	0.86	0.31	-0.02	0.04	0.00	-0.01

Note. Q_{min}, P_{mean}, P_{max} respectively refer to at-site minimum streamflow, at-site mean, and maximum daily precipitation. The rest of the variables represent percentiles of at-site daily precipitation.

Table 13. Mean of loadings and variance (in parenthesis) of all climatological variables on each principal component (PC) across all study sites.

Variables	PC1 (88%)	PC2 (10%)	PC3 (2%)	PC4 (1%)	PC5 (0%)	PC6 (0%)	PC7 (0%)	PC8 (0%)	PC9 (0%)
Q_{max}	0.19	0.88	0.40	-0.17	0.02	0.01	0.00	0.00	0.00
P_{mean}	0.32	-0.33	0.67	0.23	-0.48	0.20	-0.13	0.04	0.00
P_{70}	0.34	-0.25	0.27	-0.23	0.42	-0.61	-0.05	-0.14	-0.36
P_{75}	0.35	-0.16	0.07	-0.21	0.26	0.16	0.53	-0.11	0.64
P_{80}	0.35	-0.07	-0.14	-0.25	0.32	0.41	-0.48	0.54	-0.03
P_{90}	0.35	0.02	-0.26	-0.14	-0.18	0.37	0.44	-0.19	-0.62
P_{95}	0.35	0.05	-0.32	-0.09	-0.22	-0.03	-0.51	-0.64	0.23
P_{99}	0.35	0.07	-0.31	-0.04	-0.49	-0.50	0.15	0.49	0.14
P_{max}	0.35	0.14	-0.14	0.85	0.33	-0.02	0.03	0.00	-0.01

Note. Q_{max}, P_{mean}, P_{max} respectively refer to at-site maximum streamflow, at-site mean, and maximum daily precipitation. The rest of the variables represent percentiles of at-site daily precipitation.

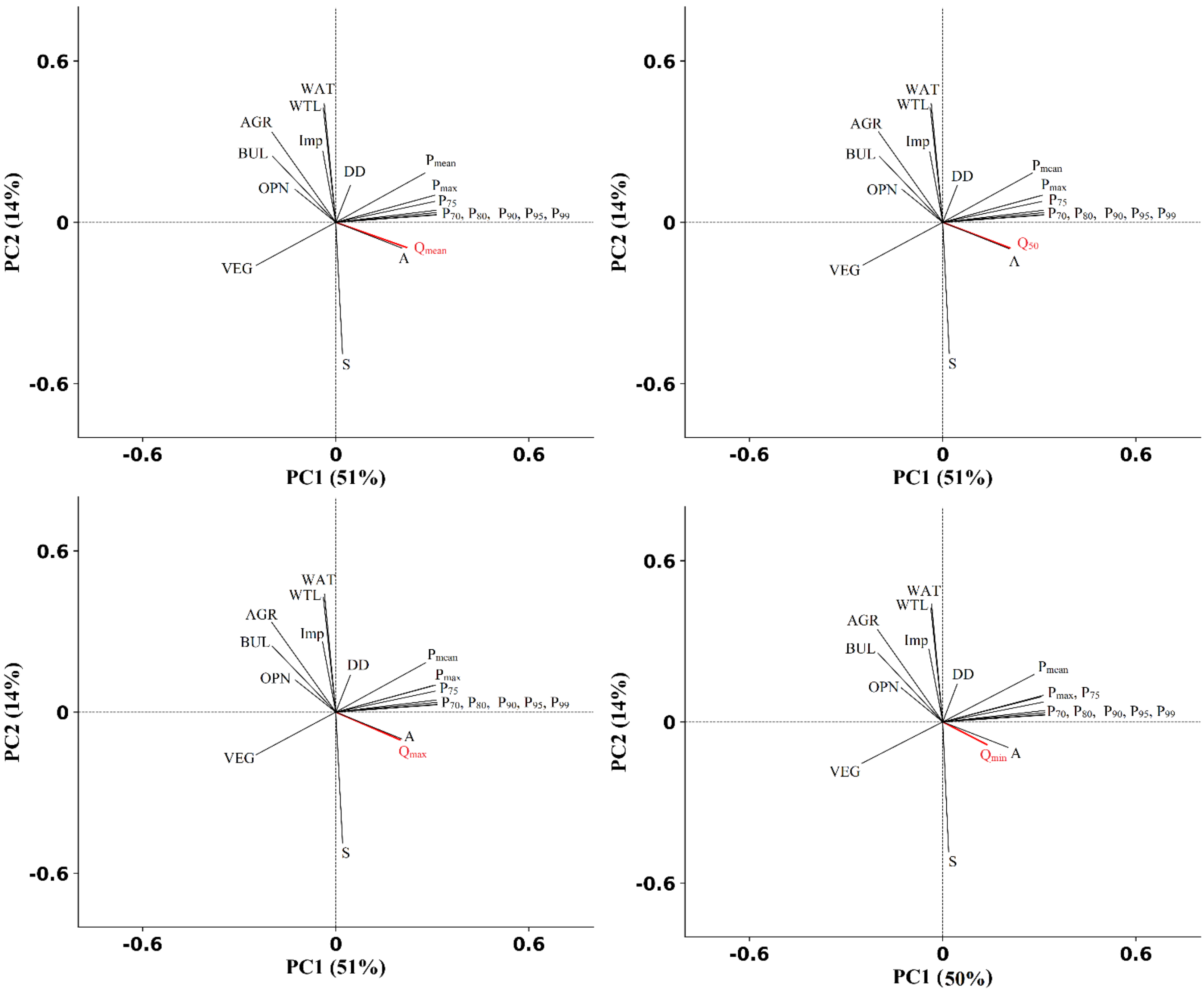

Figure 5. Biplots from principal component (PC) analysis, showing the interrelations and grouping patterns among clima-tological and land cover variables.

In the third analysis, all the hydrologic and climatological variables were associated in the study and their linkage on flow regimes was investigated (Table 33-36 in Appendix). Based on the results, the first three PCs accounted for 50%,14%, and 11% of data variance respectively. Similar to the previous results, the area had the strongest linkage with flow regimes. Following that, climatological variables had relatively strong positive linkage with flow regimes. Furthermore, land cover components implied a negative influence on streamflow regimes. Also, vegetation demonstrated a negative loading (-0.32) on PC1 in contrast to the same positive loading of mean streamflow (0.32), indicating a moderate control. Also, slope loaded substantially on PC2 (-0.43), while drainage density loaded the least on the first PC (0.02), showing their lack of control on flow regimes (Figure 5)

4.3 Emergent scaling models of streamflow regimes

Based on evaluations of the watershed's features, drainage area and mean precipitation were selected as streamflow predictors. A forward selection approach was used for adding predictors in regression. Multiple linear regression analysis yielded regression analysis results, from which the NSE and RSR are calculated (Table 15, and Table 37 in Appendix). Following a robust estimation process that involved 1,000 iterations and cross-validation, the ensemble mean of the scaling model's parameters was provided:

$$Q_{mean} = 10^{-1.62}.A^{0.99} \quad (7)$$

$$Q_{mean} = 10^{-1.71}.A^{0.93}.P_{mean}^{0.24} \quad (8)$$

$$Q_{mean} = 10^{-1.16}.A^{0.88}.(BUL)^{-0.22} \quad (9)$$

$$Q_{mean} = 10^{-1.49}.A^{0.97}.(AGR)^{-0.07} \quad (10)$$

Where $\mathrm{Q_{mean}}$, P_{mean}, A, BUL, and AGR represents mean stream flow in m^3/s, mean precipitation in mm/day, Area in km^2, built up land cover, and agricultural land cover in percentage respectively. All computed parameters were statistically significant (*p* value < 0.05) at the 95% confidence level.

The results showed that the single-predictor model using drainage area (A) as the predictor had a high level of accuracy, as measured by NSE (0.98 in calibration and 0.96 in validation). The two-predictor model with mean precipitation (P_{mean}) and A as predictors improved the accuracy of mean streamflow predictions by up to 2% compared to the single-predictor model (Figure 6, and 7). However, the two-predictor model with built up land (BUL) and A did not significantly improve the accuracy. Due to its balance between model simplicity and performance, the two-predictor model with mean precipitation (P_{mean}) and area was selected as the final model.

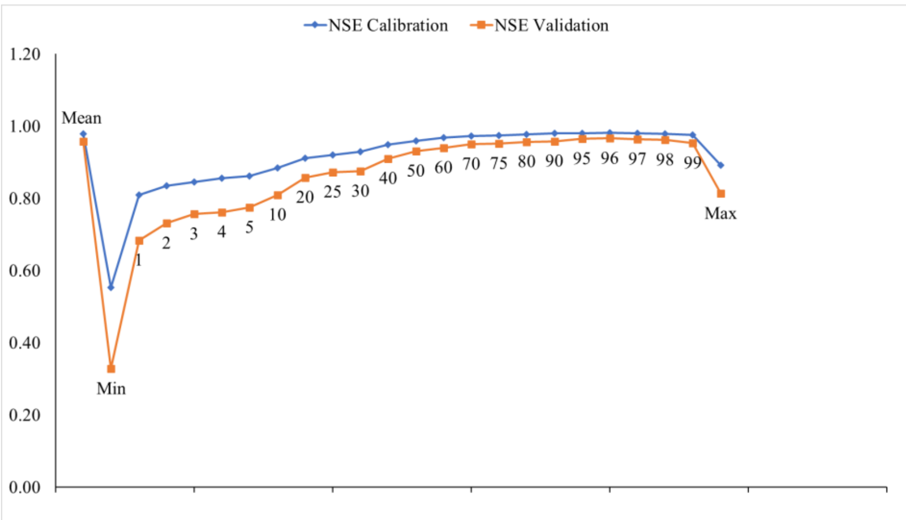

Figure 6. NSE coefficient for streamflow percentiles using Area as only predictor.

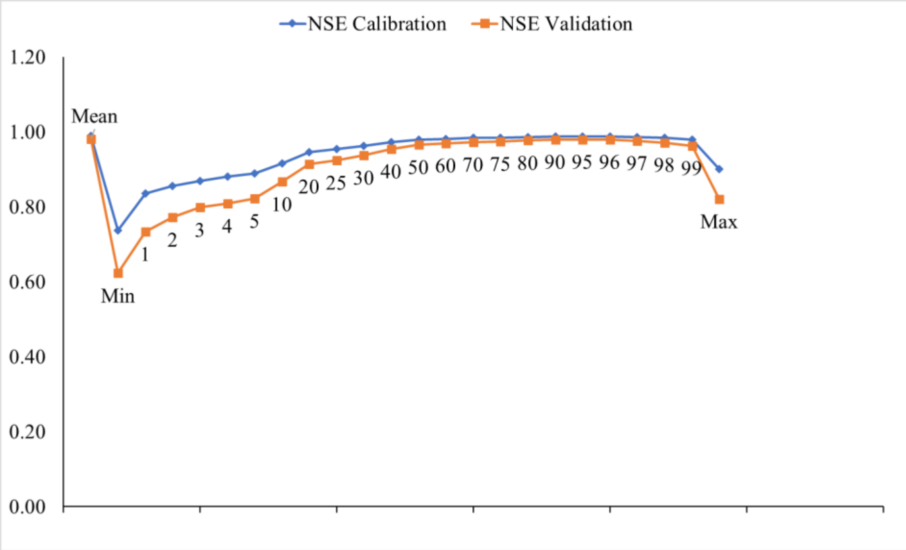

Figure 7. NSE coefficient for streamflow percentiles using Area, and mean precipitation as predictors.

The estimated scaling exponents for different flow regimes using mean precipitation and area were given (Table 15). The exponents of A were statistically significant at the 95% level of confidence (*p*-value = 0.00) for all flow regimes. Also scaling exponent of P_{mean} for Q_{mean}, Q_{10} to Q_{98} was statistically significant (*p*-value = 0.01-0.02). Whereas, the exponents of P_{mean} for Q_3, Q_4 ,Q_5 and Q_{98} flow regimes were marginally significant at the 90% level of confidence (*p*-value = 0.07- 0.12) and for the rest of the flow regimes including Q_{min}, Q_2, and Q_{max} were not statistically significant (*p*-value = 0.21-0.35).

The estimated exponents for flow regimes for the A varied from 1.37–1.13 for the Q_1-Q_{10} regime, from 1.03–0.86 for the Q_{20}-Q_{99} regime, and 1.55,0.95, and 0.8 for Q_{min}, Q_{50}, and Q_{max} respectively. The exponents of A for flow regimes followed a descending pattern from 1st to 99th flow regimes. However, the estimated exponents of P_{mean} exhibit varying patterns for the flow regimes. The estimated exponents for flow regimes for the P_{mean} varied from 0.53–0.47 for the Q_1-Q_{10} regime, from 0.53–0.12 for the Q_{20}-Q_{99} regime, and 0.51,0.33, and 0.00 for Q_{min}, Q_{50}, and Q_{max} respectively.

The scaling model's residual plot and the observed streamflow (without resampling) against the corresponding predictions were both shown (Figure 8, and 9).

According to results for the mean and median flow, the model performed well with a minimal number of outliers across all of the stations (Figure 8). However, for the minimum and maximum flows there are some over, and under predictions (Figure 9).

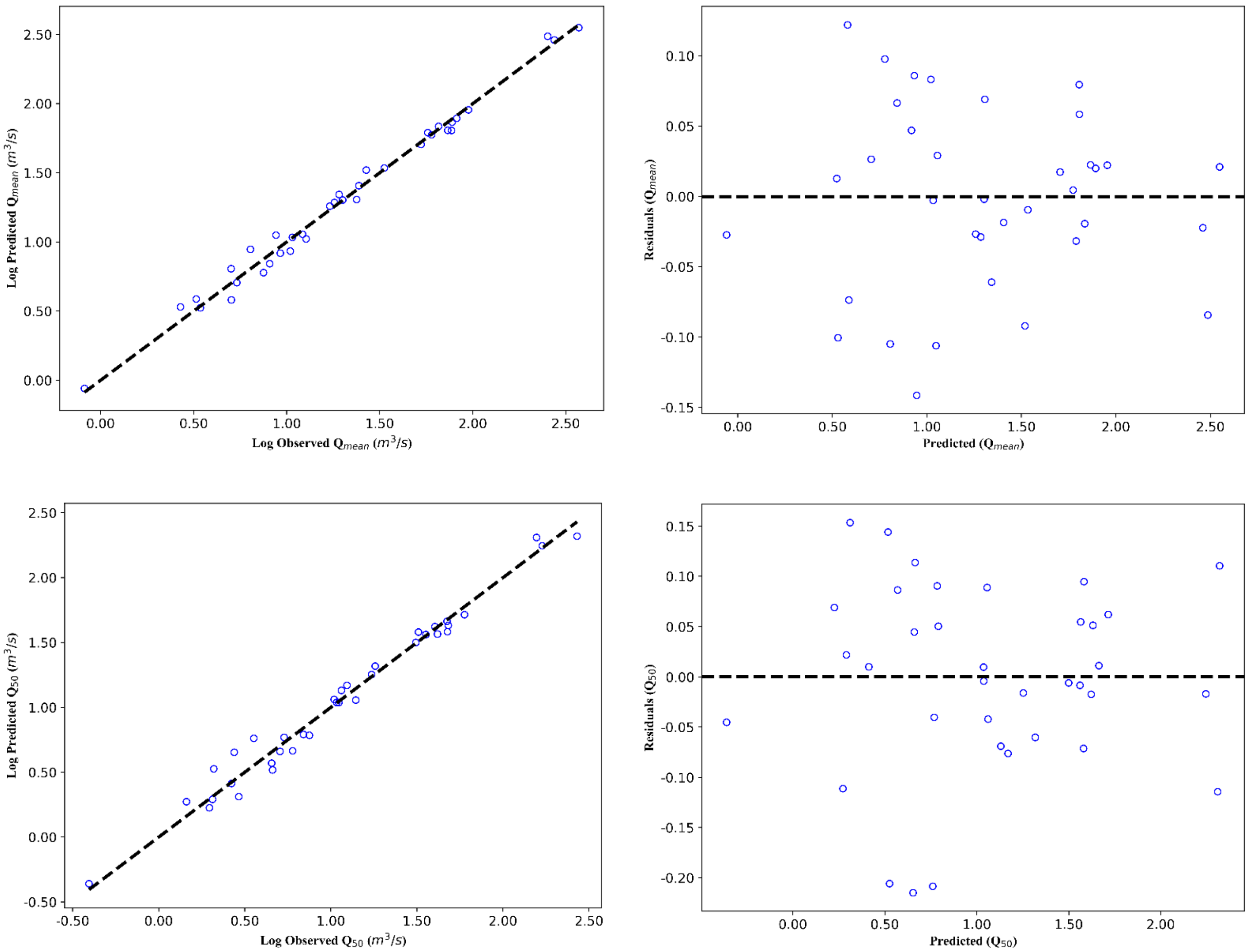

Figure 8. Observed streamflow regimes (Q) versus predicted and residual plot across all streamflow station for mean and median flow

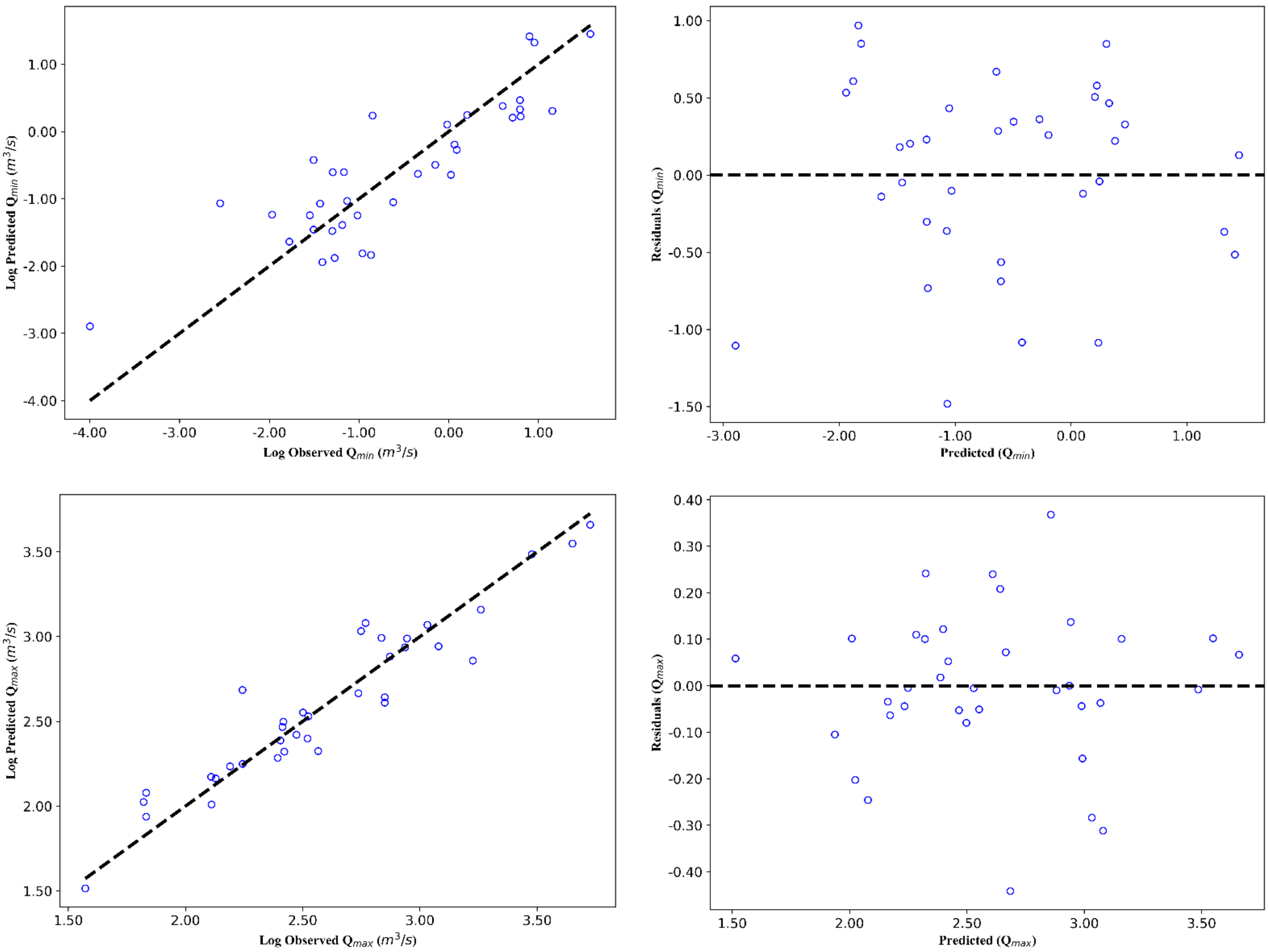

Figure 9.Observed streamflow regimes (Q) versus predicted and residual plot across all streamflow station for minimum and maximum flow

Regarding the residual plots, in each significant flow regime, they represent a consistent spread across the range of predicted values, which demonstrates the homoscedasticity of the model. This can be seen when looking at the plots. Furthermore, there is no discernible pattern or linearity in the plots that serve to validate the robust regression.

5. DISCUSSION

5.1 Dominant Controls and Processes of Flow Regimes

The evaluation of dominant controls of flow regime percentiles across the Monongahela River basin revealed that area (A) and mean precipitation (P_{mean}) are the primary factors influencing the flow regime. The size of the sub-basins played a critical role in determining the volume of water flowing through it. For minimum and maximum flow regimes, only area was dominant control. However, for other percentiles of flow regimes, the dominant controls were the same as mean streamflow. For instance, for the flow regime from 10 to 60 percentiles, area, and mean precipitation were the dominant controls.

The power-law relationships between A and P_{mean} with different flow regimes, highlighting the variation in the estimated scaling exponents. The scaling exponents for A decreased with increasing flow magnitude for all flow regimes, indicating that the scaling behavior of streamflow for A becomes weaker at higher flow magnitudes. Meanwhile, the scaling exponents for P_{mean} showed no specific pattern for Q_{min} to Q_{20} flow regimes, since the scaling exponent were not statistically significant in this range. Nonetheless, it followed the same pattern for the remaining flow regimes, and the scaling exponent reached zero at maximum flow, indicating that it had no effect on the scaling model. The scaling exponents for area are similar to the results of Gupta and Dawdy (1995) which calculated for northern mountains in Appalachia region.

The positive and significant correlation of flow regimes with basin area was consistent with other studies' conclusions (Ayalew et al., 2014; Jothityangkoon & Sivapalan, 2001; Lee & Huang, 2016). Using a nested approach to determine sub-basin area results in a robust scaling exponent with a low standard deviation when the impact of the drainage network in the observed power-law scaling statistics is considered.

The positive linkage of flow regimes with mean precipitation suggested that an increase in precipitation generally leads to an increase in streamflow. This increase in flow can be observed in different flow regime percentiles such as high flows, median flows and low flows. It is important to note that other factors can also influence the relationship between precipitation and streamflow, such as the topography and land use of the basin. Although, some of the land cover characteristics had moderate to strong correlation with flow regimes (Table 4), based on our analyses considering important characteristics that might have potential effects on flow regimes (BUL, VEG, AGR), in our study area, land use did not have robust linkage with flow regimes (Table S37 in Appendix).

5.2 Effects of Watershed land characteristics

The influence of land cover characteristics on the controlling factors of streamflow was explored. The majority of the study region's land was made up of vegetated land and agricultural lands (on average 75%, and 13% respectively; Table 14). Regression analyses were conducted for the two predominant land cover types and built-up areas. (Table S37 in Appendix). Based on the results, only agricultural land cover gave robust exponents, and model with vegetation land cover is not statistically significant at a significance level of 5%. The steep slope of the watershed justified this approach. Because the effect of vegetated land is to reduce runoff, and vegetated land works in the opposite mechanism by increasing surface runoff, in high slope areas, vegetated lands had negligible effect in reducing runoff and velocity, whereas high slope galvanizes the effect of agricultural land on streamflow (C. Yang et al., 2011).

Table 14. Average (and standard deviation in parenthesis) of percent land uses of the study watershed during 2001-2019 based on the respective NLCD data.

Land Uses	2001	2004	2006	2008	2011	2013	2016	2019
WAT (%)	0.83	0.82	0.82	0.81	0.81	0.81	0.81	0.81
	(1.15)	(1.15)	(1.12)	(1.12)	(1.14)	(1.15)	(1.11)	(1.14)
BUT (%)	8.56	8.59	8.77	8.85	9.01	9.04	9.18	9.28
	(7.34)	(7.37)	(7.68)	(7.76)	(7.93)	(7.95)	(8.03)	(8.05)
OPN (%)	0.59	0.58	0.58	0.57	0.56	0.53	0.54	0.28
	(0.63)	(0.63)	(0.63)	(0.63)	(0.63)	(0.61)	(0.60)	(0.17)
VEG (%)	75.42	75.49	75.29	75.26	75.21	75.22	75.17	75.28
	(12.60)	(12.63)	(12.80)	(12.83)	(12.90)	(12.96)	(12.99)	(13.03)
AGR (%)	13.51	13.45	13.45	13.41	13.29	13.30	13.20	13.19
	(8.15)	(8.14)	(8.12)	(8.10)	(8.03)	(8.06)	(7.99)	(8.02)
WTL (%)	0.92	0.91	0.93	0.93	0.95	0.94	0.93	0.94
	(2.33)	(2.31)	(2.36)	(2.36)	(2.39)	(2.36)	(2.36)	(2.36)

Note. WAT, BUL, OPN, VEG, AGR, WTL, Imp, A, S, DD respectively refer to water bodies, built up lands, open lands, vegetated lands, agricultural lands, wetlands, imperviousness, area, mean slope, and drainage density.

The insignificance of drainage density as an independent variable in streamflow regression is because DD had no major influence on evapotranspiration or precipitation, which were the fundamental determinants of streamflow. A higher drainage density develops faster groundwater flow yet reduces evapotranspiration. As a result, it did not have a significant impact on streamflow (Dingman, 1978; S.-Y. Yang et al., 2022). Consequently, based on results and *p*-values, built-up lands and imperviousness had a minor impact on streamflow regimes in research region, because the watershed was covered with small variation and had a low proportion of land cover across the basin.

5.3 Univariate versus Multivariate Scaling Analyses

For the univariate model (Eq (6)) area followed a linear relationship with mean streamflow. However, using approximate exponents of, respectively, 26/25 and 1/3, the emergent multivariate power law model (Equation (8)) showed a significant and scaling link of mean flow regimes with drainage area (A in Km^2) and mean precipitation(mm/day) for the Monongahela River Basin as follows:

$$Q_{mean} \propto A^{26/25}.P_{mean}{}^{1/3} \tag{11}$$

Based on the previous studies, multivariate scaling analyses helps in enhancing scaling behavior analysis (Farmer et al., 2015; Lacombe et al., 2014). However, most of the previous studies focused on several variables (7 to 13 variables) in order to improve predictive model in low flow percentiles, whereas in our model, an acceptable result was achieved (NSE= 0.62, RSR= 0.62 for Q_{min} , Table 15) with only two variables. While some models used a higher number of meteorological factors for streamflow prediction, it is crucial to highlight that the inclusion of extra variables does not necessarily guarantee better results.

Table 15. Mean and Standard Deviation (in Parenthesis) of emergent scaling factors and exponents for all flow regimes in Monongahela River Basin, USA

Flow regimes (m3/s)	Intercept	Area (Km2)	P_{mean}(mm/d)	P-value (Intercept)	P-value (Area)	P-value (P_{mean})	NSE (Val)	RSR (Val)	AIC (Val)
Q_{mean}	-1.71 (0.09)	0.93 (0.03)	0.24 (0.07)	0.00	0.00	0.00	0.98	0.13	-5.06
Q_{min}	-5.7 (0.87)	1.53 (0.28)	0.53 (0.54)	0.00	0.00	0.21	0.62	0.62	-0.65
Q_1	-4.78 (0.55)	1.37 (0.17)	0.51 (0.43)	0.00	0.00	0.12	0.73	0.51	-1.42
Q_2	-4.49 (0.48)	1.31 (0.16)	0.5 (0.39)	0.00	0.00	0.12	0.77	0.47	-1.65
Q_3	-4.26 (0.44)	1.27 (0.14)	0.46 (0.37)	0.00	0.00	0.09	0.80	0.44	-1.84
Q_4	-4.11 (0.4)	1.24 (0.13)	0.48 (0.33)	0.00	0.00	0.07	0.81	0.43	-1.99
Q_5	-4.00 (0.37)	1.21 (0.11)	0.49 (0.32)	0.00	0.00	0.06	0.82	0.42	-2.07
Q_{10}	-3.62 (0.32)	1.13 (0.10)	0.53 (0.25)	0.00	0.00	0.01	0.87	0.36	-2.50
Q_{20}	-3.08 (0.25)	1.03 (0.08)	0.53 (0.18)	0.00	0.00	0.00	0.91	0.29	-3.09
Q_{25}	-1.78 (0.11)	0.95 (0.04)	0.26 (0.09)	0.00	0.00	0.00	0.92	0.27	-3.33
Q_{30}	-2.66 (0.19)	0.97 (0.06)	0.47 (0.14)	0.00	0.00	0.00	0.94	0.24	-3.65
Q_{40}	-2.34 (0.17)	0.95 (0.05)	0.4 (0.12)	0.00	0.00	0.00	0.95	0.21	-4.05
Q_{50}	-2.11 (0.13)	0.95 (0.04)	0.33 (0.11)	0.00	0.00	0.00	0.97	0.18	-4.38
Q_{60}	-1.94 (0.12)	0.96 (0.04)	0.27 (0.09)	0.00	0.00	0.00	0.97	0.17	-4.60
Q_{70}	-1.78 (0.11)	0.95 (0.04)	0.26 (0.09)	0.00	0.00	0.00	0.97	0.16	-4.72
Q_{75}	-1.78 (0.11)	0.95 (0.04)	0.26 (0.09)	0.00	0.00	0.00	0.97	0.15	-4.77
Q_{80}	-1.62 (0.10)	0.96 (0.03)	0.24 (0.09)	0.00	0.00	0.00	0.98	0.14	-4.87
Q_{90}	-1.35 (0.09)	0.94 (0.03)	0.23 (0.08)	0.00	0.00	0.00	0.98	0.14	-5.03
Q_{95}	-1.08 (0.09)	0.91 (0.03)	0.2 (0.08)	0.00	0.00	0.00	0.98	0.14	-5.11
Q_{96}	-0.98 (0.09)	0.9 (0.03)	0.18 (0.07)	0.00	0.00	0.00	0.98	0.14	-5.09
Q_{97}	-0.87 (0.10)	0.89 (0.03)	0.16 (0.08)	0.00	0.00	0.01	0.98	0.15	-5.02
Q_{98}	-0.74 (0.10)	0.87 (0.03)	0.15 (0.08)	0.00	0.00	0.02	0.97	0.16	-4.91
Q_{99}	-0.56 (0.11)	0.86 (0.04)	0.12 (0.09)	0.00	0.00	0.08	0.96	0.18	-4.69
Q_{max}	0.26 (0.19)	0.79 (0.07)	0.01 (0.19)	0.10	0.00	0.35	0.82	0.41	-3.29

Moreover, models incorporating additional variables and incorporating land cover characteristics in the scaling study were analyzed, indicating that their inclusion did not result in significant improvement of the model. As a result, there was no evidence of omitted variable bias or inconsistency in the calculation of scaling exponents. In fact, the simpler models presented here, which just consisted of essential factors including area and precipitation, produced the same precise predictions as more complicated models. A simpler model can also make the modeling process more accessible and efficient by lowering the number of variables.

The variation in the estimated exponents of A and P suggested the multiscaling property of streamflow in the Monongahela River Basin. This behavior was dominant in mean, and 20^{th} to 96^{th} of flow regimes. Based on the results, area had the strong almost linear impact on streamflow. This finding in our study area had not been investigated and is in line with the finding in other study areas (Gupta et al., 1994; Gupta & Dawdy, 1995; Zhang et al., 2020). According to their statement, the area exponent ranges between 0.65 to 1.0 and tends to rise as more variables are used in a multivariate regression. Furthermore, based on our findings and past research, the scaling exponent appeared to become more linear as basin area increases. This suggests that in the large basins, the relationship between the watershed area and flow qualities becomes more predictable(Goodrich et al., 1997). It was noteworthy that the model with area as the only exponent was valid in all flow regimes, but the two variable models were statistically significant valid for just part of flow regimes (10^{th} to 97^{th} and mean streamflow). Yet, under robust flow regimes, the model with climatological parameter and area outperformed the model with area only. The graphical representation of scaling laws and organizing principles was achieved by plotting critical streamflow percentiles (calculated using Equation (8)) against drainage area for different mean precipitation values (as shown in Figure 4).

Overall, by considering both area and precipitation, our multivariate model is able to better capture the interactions between these two important factors, leading to more accurate and reliable streamflow predictions especially in extremum flows. In contrast, a univariate model based solely on catchment area may fail to adequately account for the temporal and spatial variability of rainfall events, thus rendering it less capable of accurately predicting streamflow.

5.4 Generality, Applications, and Limitations of the Scaling Law

Application of scaling models is to estimate streamflow in ungauged basins, where there are no streamflow observations. This is particularly useful for water resources planning and management, where it is important to know the streamflow that is available in a basin for different uses, such as irrigation, domestic water supply, and hydropower generation. Scaling models were utilized to predict streamflow regimes at the site level and demonstrated strong predictive performance across the entire watershed area. (NSE = 0.62 to 0.98, and RSR = 0.15 to 0.62).

Streamflow scaling models that incorporate watershed area and precipitation could provide several benefits in predicting streamflow in gauged and ungauged basins. Firstly, the watershed area was an important factor in determining streamflow as it influences the water that could be stored and transmitted through the watershed. Streamflow scaling models that incorporate watershed areas could account for differences in the size of watersheds, which could vary widely across different basins (Figure 10). This could help to improve the accuracy of streamflow predictions, particularly in large watersheds where the impact of the basin size on streamflow are significant (Vogel et al., 1999).

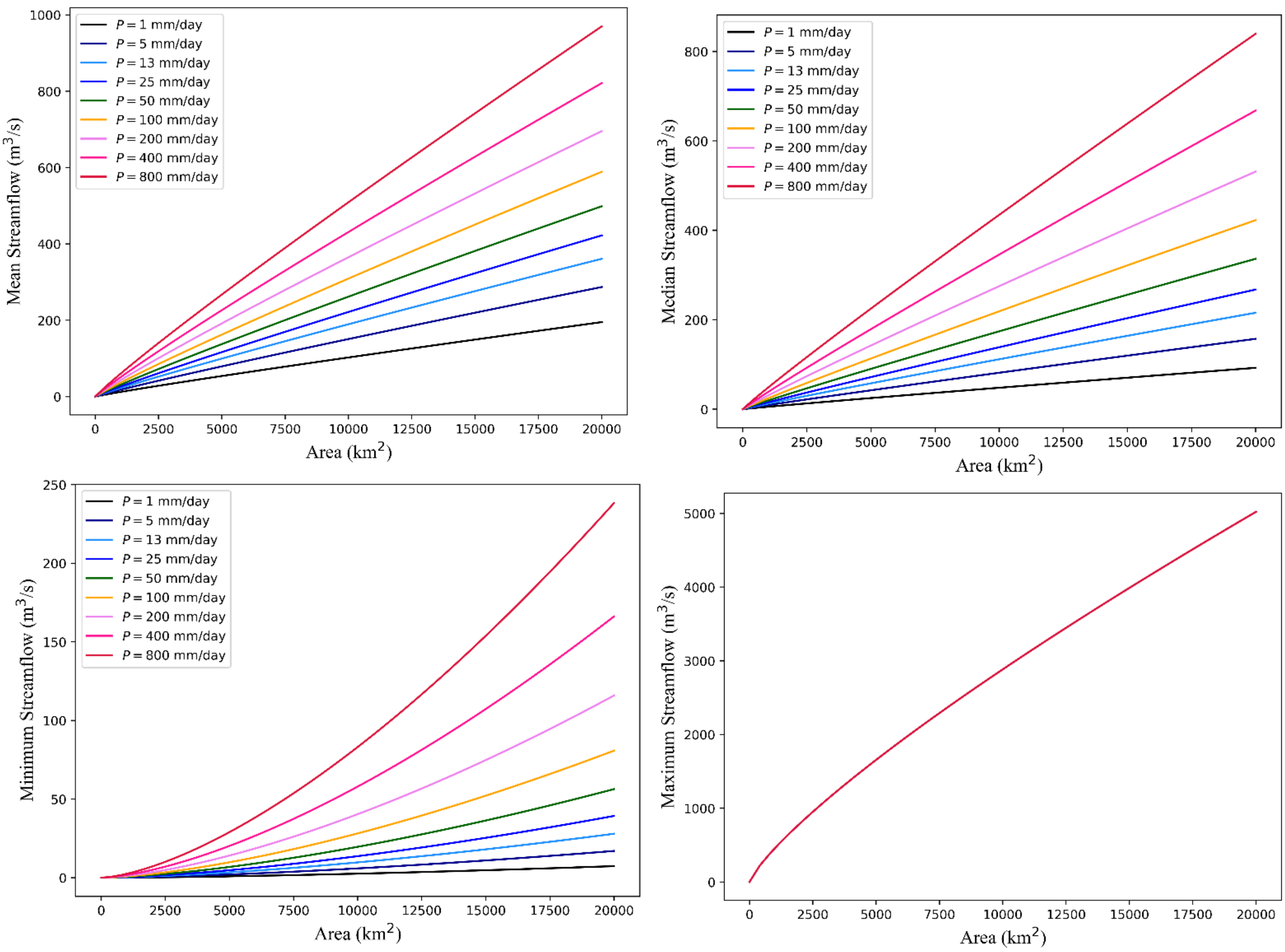

Figure 10.Plot of streamflow percentiles versus drainage area (A) for different mean precipitation values. streamflow was estimated using the emergent power law scaling relation of Equation 8

Secondly, precipitation was a key driver of streamflow, and streamflow scaling models that include precipitation captured the influence of rainfall and snowmelt on streamflow (Khan et al., 2022). It improved streamflow predictions in areas where precipitation patterns vary significantly, such as regions with highly variable rainfall or snowfall amounts. By incorporating precipitation data, streamflow scaling models accounted for the impact of both short-term and long-term changes in precipitation patterns on streamflow (Ayalew et al., 2014).

Furthermore, incorporating watershed characteristics in streamflow scaling models also provided a useful framework for understanding the hydrologic response of a watershed. Future research can help to identify the dominant hydrological processes in the watershed, such as the role of groundwater storage, and can provide insights into the factors that control streamflow in different basins.

6. CONCLUSION

This study provided a comprehensive assessment of streamflow scaling in the large mountainous Monongahela River Basin. The investigation focused on the scaling of flow regimes, with a particular emphasis on the impact of climatological and watershed characteristics. Area and mean precipitation were discovered to be the most important factors influencing streamflow percentiles over a 30-year study period.

In order to study the scaling of streamflow, a nested approach was employed. This approach involved dividing the basin into smaller sub-basins and analyzing the streamflow data for each sub-basin separately. Through this analysis, it was discovered that the types of land cover in the basin, such as vegetated and agricultural land, had little to no impact on streamflow scaling. This was somewhat unexpected, as these types of land cover were dominant throughout the basin.

Several preprocessing steps were taken to prepare the streamflow data for analysis. Pearson correlation was utilized to identify any patterns or relationships between different variables that may be affecting the streamflow. Additionally, principal component analyses were conducted to identify any significant factors that may be influencing streamflow. These steps were crucial in ensuring that the streamflow data was properly analyzed.

The study found that the scaling properties of all streamflow percentiles followed a power law scaling function with multiscaling properties. This means that the relationship between the streamflow percentiles and their dominant variables could be described by a power law function, and that these properties varied across different percentiles.

The results of this study were consistent with those of other principal studies, which have shown that the area exponent tends to increase with the number of watersheds and climatic variables used in a multivariate regression analysis. This indicates that the slopes of the power law function vary systematically with flow regimes in the scaling investigation of streamflow. In other words, the scaling properties of the streamflow vary depending on the flow regime and the characteristics of the watershed.

Understanding the scaling properties of streamflow is important for predicting and managing flood events, as it allows us to better understand how streamflow behaves across different scales. By analyzing the scaling properties of streamflow, we can identify the relationships between streamflow and other hydrological factors and use this information to develop more accurate models for predicting and managing flood events.

One important application of the scaling laws is the prediction of streamflow in ungauged basins. Ungauged basins are those that do not have streamflow data readily available, and their prediction is essential for effective water resources management. The scaling laws were found to yield reliable outcomes for streamflow prediction across watersheds in ungauged basins, with NSE values ranging from 0.55 to 0.98 and RSR values ranging from 0.15 to 0.70.

To predict streamflow in ungauged basins, the model was trained on 70% of the available data and then simulated the remaining 30% as if it represented ungauged basins omitted from the regression analysis. The researchers were able to evaluate the performance of the scaling laws for predicting streamflow in ungauged basins using this method. The results showed that the scaling laws could predict streamflow accurately in ungauged basins, with high NSE and low RSR values indicating good agreement between predicted and observed streamflow.

The research's results were first put into practice within the Monongahela watershed area, and they will aid water resource scientists and managers in getting a better understanding of the watershed's hydrologic response and insights into the main hydrological processes regulating streamflow in various basins. One of the strengths of this study is the use of a nested approach to analyze the scaling properties of streamflow. This approach allowed for a more detailed and nuanced analysis of the relationships between streamflow and other hydrological factors at different scales.

Future research in this area could concentrate on expanding the analysis to other watersheds and geographical regions to assess the generalizability of the findings. The study was limited to a single geographic area and may not be applicable to other areas. Furthermore, future research could investigate the use of various modeling approaches and techniques to analyze the scaling properties of streamflow and their application in predicting streamflow in ungauged basins. Machine learning algorithms and artificial intelligence techniques, for example, could provide a more accurate and efficient way to predict streamflow in ungauged basins.

In summary, this study provides insights into the scaling properties of streamflow in mountainous watersheds with mixed landcovers. It highlights the potential for using scaling properties to predict streamflow in ungauged areas, which could have important implications for water resource management and planning.

7. REFERENCES

Abdul-Aziz, O. I., & Ahmed, S. (2017). Relative linkages of stream water quality and environmental health with the land use and hydrologic drivers in the coastal-urban watersheds of southeast Florida. *GeoHealth*, *1*(4), 180–195. https://doi.org/10.1002/2017GH000058

Abdul-Aziz, O. I., & Gebreslase, A. K. (2023). Emergent Scaling of Dissolved Oxygen (DO) in Freshwater Streams Across Contiguous USA. *Water Resources Research*, *59*(2), e2022WR032114. https://doi.org/10.1029/2022WR032114

Abdul-Aziz, O. I., Ishtiaq, K. S., Tang, J., Moseman-Valtierra, S., Kroeger, K. D., Gonneea, M. E., Mora, J., & Morkeski, K. (2018). Environmental Controls, Emergent Scaling, and Predictions of Greenhouse Gas (GHG) Fluxes in Coastal Salt Marshes. *Journal of Geophysical Research: Biogeosciences*, *123*(7), 2234–2256. https://doi.org/10.1029/2018JG004556

Abdul Razaq, S., Tarmizi, I., Heryansyah, A., Lawan, U. F., Alamgir, M., & Hadi Pour, S. (2016). Streamflow prediction in ungauged catchments in the east coast of Peninsular Malaysia using multivariate statistical techniques. *Jurnal Teknologi*, *78*(6–12), 43–49. https://doi.org/10.11113/jt.v78.9231

Adams, T. E., & Pagano, T. C. (2016). Chapter 9 - Flood Forecasting — A National Overview for Great Britain. In *Flood Forecasting: A Global Perspective*. Elsevier Inc. https://doi.org/10.1016/C2014-0-01361-5

Al-Rawas, G. A., & Valeo, C. (2010). Relationship between wadi drainage characteristics and peak-flood flows in arid northern Oman. *Hydrological Sciences Journal*, *55*(3), 377–393. https://doi.org/10.1080/02626661003718318

Alfieri, L., Salamon, P., Pappenberger, F., Wetterhall, F., & Thielen, J. (2012). Operational early warning systems for water-related hazards in Europe. *Environmental Science & Policy*, *21*, 35–49. https://doi.org/10.1016/J.ENVSCI.2012.01.008

Arnold, C. L., & Gibbons, & C. J. (1996). Impervious Surface Coverage: The Emergence of a Key Environmental Indicator. *Journal of the American Planning Association*, *62*(2), 243–258. https://doi.org/10.1080/01944369608975688

Ayalew, T. B., Krajewski, W. F., & Mantilla, R. (2014). Connecting the power-law scaling structure of peak-discharges to spatially variable rainfall and catchment physical properties. *Advances in Water Resources*, *71*, 32–43. https://doi.org/10.1016/J.ADVWATRES.2014.05.009

Ayalew, T. B., Krajewski, W. F., & Mantilla, R. (2015). Analyzing the effects of excess rainfall properties on the scaling structure of peak discharges: Insights from a mesoscale river basin. *Water Resources Research*, *51*(6), 3900–3921. https://doi.org/10.1002/2014WR016258

Barnett, T. P., Adam, J. C., & Lettenmaier, D. P. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature 2005 438:7066*, *438*(7066), 303–309. https://doi.org/10.1038/nature04141

Benson, M. A. (1964). Factors affecting the occurrence of floods in the southwest. *U.S.GEOLOGICAL SURVEY WATER-SUPPLY PAPER*, *1580-D*, 72.

Berghuijs, W. R., Woods, R. A., & Hrachowitz, M. (2014). A precipitation shift from snow towards rain leads to a decrease in streamflow. *Nature Climate Change 2014 4:7*, *4*(7), 583–586. https://doi.org/10.1038/nclimate2246

Blöschl, G., & Sivapalan, M. (1995). Scale issues in hydrological modelling: A review. *Hydrological Processes*, *9*(3–4), 251–290. https://doi.org/10.1002/HYP.3360090305

Bosch, J. M., & Hewlett, J. D. (1982). A review of catchment experiments to determine the effect of vegetation changes on water yield and evapotranspiration. *Journal of Hydrology*, *55*(1–4), 3–23. https://doi.org/10.1016/0022-1694(82)90117-2

Bourdin, D. R., Fleming, S. W., & Stull, R. B. (2012). Streamflow modelling: A primer on applications, approaches and challenges. *Atmosphere - Ocean*, *50*(4), 507–536. https://doi.org/10.1080/07055900.2012.734276

Brassel, K. E., & Reif, D. (1979). A Procedure to Generate Thiessen Polygons. *Geographical Analysis*, *11*(3), 289–303. https://doi.org/10.1111/J.1538-4632.1979.TB00695.X

Buttle, J. M., & Eimers, M. C. (2009). Scaling and physiographic controls on streamflow behaviour on the Precambrian Shield, south-central Ontario. *Journal of Hydrology*, *374*(3–4), 360–372. https://doi.org/10.1016/j.jhydrol.2009.06.036

Deems, J. S., Fassnacht, S. R., & Elder, K. J. (2006). Fractal Distribution of Snow Depth from Lidar Data. *Journal of Hydrometeorology*, *7*(2), 285–297. https://doi.org/10.1175/JHM487.1

Dingman, S. L. (1978). Drainage density and streamflow: A closer look. *Water Resources Research*, *14*(6), 1183–1187. https://doi.org/10.1029/WR014i006p01183

Dunne, T., & Leopold, L. (1958). Water in Environmental Planning. *Progress in Physical Geography: Earth and Environment, 4(2), 301–304*, *60*, 1–6. https://doi.org/https://doi.org/10.1177/030913338000400216

Ellenburg, W. L., Cruise, J. F., & Singh, V. P. (2018). The role of evapotranspiration in streamflow modeling – An analysis using entropy. *Journal of Hydrology*, *567*, 290–304. https://doi.org/10.1016/J.JHYDROL.2018.09.048

Farmer, W. H., Over, T. M., & Vogel, R. M. (2015). Multiple regression and inverse moments improve the characterization of the spatial scaling behavior of daily streamflows in the Southeast United States. *Water Resources Research*, *51*(3), 1775–1796. https://doi.org/10.1002/2014WR015924

Formetta, G., Over, T., & Stewart, E. (2021). Assessment of Peak Flow Scaling and Its Effect on Flood Quantile Estimation in the United Kingdom. *Water Resources Research*, *57*(4). https://doi.org/10.1029/2020WR028076

Fraedrich, K., & Larnder, C. (2016). Scaling regimes of composite rainfall time series. *New Pub: Stockholm Uni Press*, *45 A*(4), 289–298. https://doi.org/10.3402/TELLUSA.V45I4.14893

Goodrich, D. C., Lane, L. J., Shillito, R. M., Miller, S. N., Syed, K. H., & Woolhiser, D. A. (1997). Linearity of basin response as a function of scale in a semiarid watershed. *Water Resources Research*, *33*(12), 2951–2965. https://doi.org/10.1029/97WR01422

Gupta, V. K. (2004). Emergence of statistical scaling in floods on channel networks from complex runoff dynamics. *Chaos, Solitons & Fractals*, *19*(2), 357–365. https://doi.org/10.1016/S0960-0779(03)00048-1

Gupta, V. K., & Dawdy, D. R. (1995). Physical interpretations of regional variations in the scaling exponents of flood quantiles. *Hydrological Processes*, *9*(3–4), 347–361. https://doi.org/10.1002/HYP.3360090309

Gupta, V. K., Mesa, O. J., & Dawdy, D. R. (1994). Multiscaling theory of flood peaks: Regional quantile analysis. *Water Resources Research*, *30*(12), 3405–3421. https://doi.org/10.1029/94WR01791

Gupta, V. K., & Waymire, E. (1990). Multiscaling properties of spatial rainfall and river flow distributions. *Journal of Geophysical Research*, *95*(D3), 1999–2009.

https://doi.org/10.1029/JD095iD03p01999

Hack, J. T. (1957). Studies of longitudinal stream profiles in Virginia and Maryland. In *USGS Professional Paper 249*.

Horton. (1945). *EROSIONAL DEVELOPMENT OF STREAMS AND THEIR DRAINAGE BASINS; HYDROPHYSICAL APPROACH TO QUANTITATIVE MORPHOLOGY*. GeoScienceWorld. https://doi.org/10.1130/0016-7606(1945)56[275:EDOSAT]2.0.CO;2

Ishak, E., Haddad, K., Zaman, M., & Rahman, A. (2011). Scaling property of regional floods in New South Wales Australia. *Natural Hazards*, *58*(3), 1155–1167. https://doi.org/10.1007/S11069-011-9719-6

Jolliffe, I. T. (2002). Principal Component Analysis. *Ieee Signal Processing Letters*, *9*(2), 40–42. https://doi.org/10.1007/B98835

Jothityangkoon, C., & Sivapalan, M. (2001). Temporal scales of rainfall–runoff processes and spatial scaling of flood peaks: space–time connection through catchment water balance. *Advances in Water Resources*, *24*(9–10), 1015–1036. https://doi.org/10.1016/S0309-1708(01)00044-6

Khan, Z., Khan, F. A., Khan, A. U., Hussain, I., Khan, A., Shah, L. A., Khan, J., Badrashi, Y. I., Kamiński, P., Dyczko, A., & Różkowski, K. (2022). Climate-Streamflow Relationship and Consequences of Its Instability in Large Rivers of Pakistan: An Elasticity Perspective. *Water 2022, Vol. 14, Page 2033*, *14*(13), 2033. https://doi.org/10.3390/W14132033

Kirkby, M. (1988). Hillslope runoff processes and models. *Journal of Hydrology*, *100*(1–3), 315–339. https://doi.org/10.1016/0022-1694(88)90190-4

Kroll, C. N., Rapant, D. F., & Vogel, R. M. (2014). Prediction of Hydrologic Statistics in Nested Watersheds across the United States. *World Environmental and Water Resources Congress 2014: Water Without Borders - Proceedings of the 2014 World Environmental and Water Resources Congress*, 2326–2335. https://doi.org/10.1061/9780784413548.232

Kuhn, M., & Johnson, K. (2013a). An Introduction to Feature Selection. In: Applied Predictive Modeling. . *Springer, New York, NY*, 487–519. https://doi.org/10.1007/978-1-4614-6849-3_19

Kuhn, M., & Johnson, K. (2013b). Linear Regression and Its Cousins. In: Applied Predictive Modeling. *Springer, New York, NY*, 101–139. https://doi.org/10.1007/978-1-4614-6849-3_6

Lacombe, G., Douangsavanh, S., Vogel, R. M., McCartney, M., Chemin, Y., Rebelo, L. M., & Sotoukee, T. (2014). Multivariate power-law models for streamflow prediction in the Mekong Basin. *Journal of Hydrology: Regional Studies*, *2*, 35–48. https://doi.org/10.1016/j.ejrh.2014.08.002

Lee, K. T., & Huang, J. K. (2016). Influence of storm magnitude and watershed size on runoff nonlinearity. *Journal of Earth System Science*, *125*(4), 777–794. https://doi.org/10.1007/S12040-016-0700-3/METRICS

Leopold, L. B., & Jr., T. M. (1953). The hydraulic geometry of stream channels and some physiographic implications. *Professional Paper*. https://doi.org/10.3133/PP252

Leopold, L. B., Wolman, M. G., & Miller, J. P. (1964). *Fluvial processes in geomorphology*. Courier Corporation.

Mcguirk, A., & Spanos, A. (2002). The Linear Regression Model with Autocorrelated Errors: Just Say No to Error Autocorrelation. *Annual Meeting of the American Agricultural Economics Association*, *1993*, 1–14.

Medhi, H., & Tripathi, S. (2015). On identifying relationships between the flood scaling exponent and basin attributes. *Chaos*, *25*(7). https://doi.org/10.1063/1.4916378

Modarres, R. (2010). Low Flow Scaling with Respect to Drainage Area and Precipitation in Northern Iran. *Journal of Hydrologic Engineering*, *15*(3), 210–214. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000181

Moriasi, D. N., Arnold, J. G., Liew, M. W. Van, Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. *Transactions of the ASABE*, *50*(3), 885–900. https://doi.org/10.13031/2013.23153

MRLC. (2022). *Multi-Resolution Land Characteristics (MRLC) Consortium*. Available Online at Https://Www.Mrlc.Gov/Data-Services-Page. (Accessed 17 November 2022). https://www.mrlc.gov/data-services-page

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I - A discussion of principles. *Journal of Hydrology*, *10*(3), 282–290. https://doi.org/10.1016/0022-1694(70)90255-6

NOAA. (2022). *National Oceanic and Atmospheric Administration*. Climate Data Online. Available Online at Https://Www.Ncei.Noaa.Gov/Cdo-Web/ (Last Accessed in May 17, 2022). . https://www.ncei.noaa.gov/cdo-web/

Özger, M., & Singh, V. P. (2013). Scaling properties of Euphrates and Tigris basin daily streamflow data EU-India River Bank Filtration Project View project. *Conferance: 6th International Perspective on Water Resources & the Environment (IPWE 2013) (ASCE-EWRI)*.

Pandey, G. R., & Nguyen, V. T. V. (1999). A comparative study of regression based methods in regional flood frequency analysis. *Journal of Hydrology*, *225*(1–2), 92–101. https://doi.org/10.1016/S0022-1694(99)00135-3

Parzen, E., & Priestley, M. B. (1984). Spectral Analysis and Time Series. *Journal of the American Statistical Association*, *79*(385), 233. https://doi.org/10.2307/2288374

Perez, G., Mantilla, R., & Krajewski, W. F. (2018). Spatial Patterns of Peak Flow Quantiles Based on Power-Law Scaling in the Mississippi River Basin. *Advances in Nonlinear Geosciences*, 497–518. https://doi.org/10.1007/978-3-319-58895-7_23

Piadeh, F., Behzadian, K., & Alani, A. M. (2022). A critical review of real-time modelling of flood forecasting in urban drainage systems. In *Journal of Hydrology* (Vol. 607). Elsevier B.V. https://doi.org/10.1016/j.jhydrol.2022.127476

Sauquet, E., Ramos, M. H., Chapel, L., & Bernardara, P. (2008). Streamflow scaling properties: Investigating characteristic scales from different statistical approaches. *Hydrological Processes*, *22*(17), 3462–3475. https://doi.org/10.1002/hyp.6952

Sharma, P. J., Patel, P. L., & Jothiprakash, V. (2021). Data-driven modelling framework for streamflow prediction in a physio-climatically heterogeneous river basin. *Soft Computing*, *25*(8), 5951–5978. https://doi.org/10.1007/S00500-021-05585-9/METRICS

Shen, M., & Chui, T. F. M. (2021). Contrasting scaling relationships of extreme precipitation and streamflow to temperature across the United States. *Environmental Research Communications*, *3*(12), 125008. https://doi.org/10.1088/2515-7620/AC40EF

Sivapalan, M., Takeuchi, K., Franks, S. W., Gupta, V. K., Karambiri, H., Lakshmi, V., Liang, X., McDonnell, J. J., Mendiondo, E. M., O’Connell, P. E., Oki, T., Pomeroy, J. W., Schertzer, D., Uhlenbrook, S., & Zehe, E. (2003). IAHS Decade on Predictions in Ungauged Basins (PUB), 2003-2012: Shaping an exciting future for the hydrological sciences. *Hydrological Sciences Journal*, *48*(6), 857–880. https://doi.org/10.1623/HYSJ.48.6.857.51421

Smith, J. A. (1992). Representation of basin scale in flood peak distributions. *Water Resources Research*, *28*(11), 2993–2999. https://doi.org/10.1029/92WR01718

Solomatine, D. P., & Ostfeld, A. (2008). Data-driven modelling: Some past experiences and new approaches. *Journal of Hydroinformatics*, *10*(1), 3–22. https://doi.org/10.2166/HYDRO.2008.015

Telesca, L., Lovallo, M., Lopez-Moreno, I., & Vicente-Serrano, S. (2012). Investigation of scaling properties in monthly streamflow and Standardized Streamflow Index (SSI) time series in the Ebro basin (Spain). *Physica A: Statistical Mechanics and Its Applications*, *391*(4), 1662–1678. https://doi.org/10.1016/j.physa.2011.10.023

Tessier, Y., Lovejoy, S., Hubert, P., Schertzer, D., & Pecknold, S. (1996). Multifractal analysis and modeling of rainfall and river flows and scaling, causal transfer functions. *Journal of Geophysical Research Atmospheres*, *101*(21), 26427–26440. https://doi.org/10.1029/96jd01799

Thomas, N. W., Ayalew, T. B., Arenas, A. A., Schilling, K. E., Weber, L. J., & Krajewski, W. F. (2021). Characterizing the effects of dry antecedent soil moisture conditions, channel transmission losses, and variable precipitation on peak flow scaling. *Advances in Water Resources*, *158*, 104061. https://doi.org/10.1016/J.ADVWATRES.2021.104061

US Army Corps of Engineers, U. (2011). Monongahela River Watershed Initial Watershed Assessment September 2011. *Assessment*, *September*, 73.

USGS. (2022). *United States geological survey*. National Water Information System. Available Online at Https://Maps.Waterdata.Usgs.Gov/Mapper/Index.Html. (Accessed 17 November 2022). https://maps.waterdata.usgs.gov/mapper/index.html

Villarini, G., Smith, J. A., Baeck, M. L., & Krajewski, W. F. (2011). Examining Flood Frequency Distributions in the Midwest U.S. *Journal of the American Water Resources Association*, *47*(3), 447–463. https://doi.org/10.1111/J.1752-1688.2011.00540.X

Vogel, R. M., & Sankarasubramanian, A. (2000). Spatial scaling properties of annual streamflow in the United States. *Hydrological Sciences-Journal-Des Sciences Hydrologiques*, *45*(3), 455. https://doi.org/10.1080/02626660009492342

Vogel, R. M., Wilson, I., & Daly, C. (1999). Regional Regression Models of Annual Streamflow for the United States. *Journal of Irrigation and Drainage Engineering*, *125*(3), 148–157. https://doi.org/10.1061/(ASCE)0733-9437(1999)125:3(148)

Yang, C., Yu, Z., Hao, Z., Lin, Z., & Wang, H. (2011). Effects of Vegetation Cover on Hydrological Processes in a Large Region: Huaihe River Basin, China. *Journal of Hydrologic Engineering*, *18*(11), 1477–1483. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000440

Yang, S.-Y., Chang, C.-H., Hsu, C.-T., Shiang, ·, & Wu, J. (2022). Variation of uncertainty of drainage density in flood hazard mapping assessment with coupled 1D-2D hydrodynamics model. *Natural Hazards*, *111*, 2297–2315. https://doi.org/10.1007/s11069-021-05138-1

Yue, S., & Yew Gan, T. (2004). Simple scaling properties of Canadian annual average streamflow. *Advances in Water Resources*, *27*(5), 481–495. https://doi.org/10.1016/j.advwatres.2004.02.019

Zhang, Z., Balay, J. W., & Liu, C. (2020). Regional regression models for estimating monthly streamflows. *Science of the Total Environment*, *706*. https://doi.org/10.1016/J.SCITOTENV.2019.135729

Zheng, Y., Li, J., Zhang, T., Rong, Y., & Feng, P. (2021). Exploring the Application of Flood Scaling Property in Hydrological Model Calibration. *Journal of Hydrometeorology*, *22*(12), 3255–3274. https://doi.org/10.1175/JHM-D-21-0123.1

APPENDIX

Text S1. Nested approach for creating cumulative data

This approach considers the watershed as a system of nested sub-systems, each with its own unique physical, characteristics, and each influencing and being influenced by the others. By considering the watershed as a system of nested sub-systems, the nested approach provided a holistic, integrated approach to watershed management (Ayalew et al., 2014). The technique for calculating nested subwatesheds was to trace the drainage network from downstream to upstream. We accumulated subbasin regions at lower stream hierarchy by following flow direction. At the top of the hierarchy, a large-scale model was utilized to simulate streamflow throughout the entire watershed basin. We created a watershed hierarchy matrix to show the rank of each subwatershed and calculated cumulative data for analyses. The nested modeling approach is useful for modeling large-scale streamflow dynamics while also providing detailed information about streamflow patterns at smaller sizes. It can be used to examine how climate change, land use changes, and other variables affect streamflow patterns (Gupta & Dawdy, 1995).

Text S2. Assigning Precipitation station to sub watershed.

The Thiessen Polygon method is a widely used spatial interpolation technique for allocating point data to a continuous surface. We used this method to assign precipitation data from 45 climate stations to 36 sub-watersheds within the study area. Firstly, we generated Thiessen polygons for each climate station, which defined the catchment area around each station. Next, we delineated the sub-watersheds using a set of watershed delineation steps in ArcGIS 2.5 and overlaid the Thiessen polygons onto the sub-watershed map to determine the precipitation stations that fell within the boundary of each sub-watershed. We then calculated the area-weighted average precipitation for each sub-watershed by multiplying the precipitation value of each station by the area of the corresponding Thiessen polygon and dividing the sum of these products by the total area of the sub-watershed (Eq. 1S). Finally, we assigned the calculated precipitation value to the centroid of each sub-watershed as the representative precipitation value for that sub-watershed. After assigning precipitation data, we used the same nested approach to generate cumulative precipitation data for data analyses. This method allowed us to interpolate precipitation data for each sub-watershed and generate a continuous precipitation surface, which can be used for hydrological modeling and water resources management purposes.

$$P_{sub} = \frac{\Sum(\mathrm{P_i} * \mathrm{A_i})}{\Sum(\mathrm{A_i})} \qquad (1S)$$

Where: P_{sub} is the area-weighted average precipitation for the sub-watershed, P_i is the precipitation value of climate station i, A_i is the area of the Thiessen polygon that covers station i.

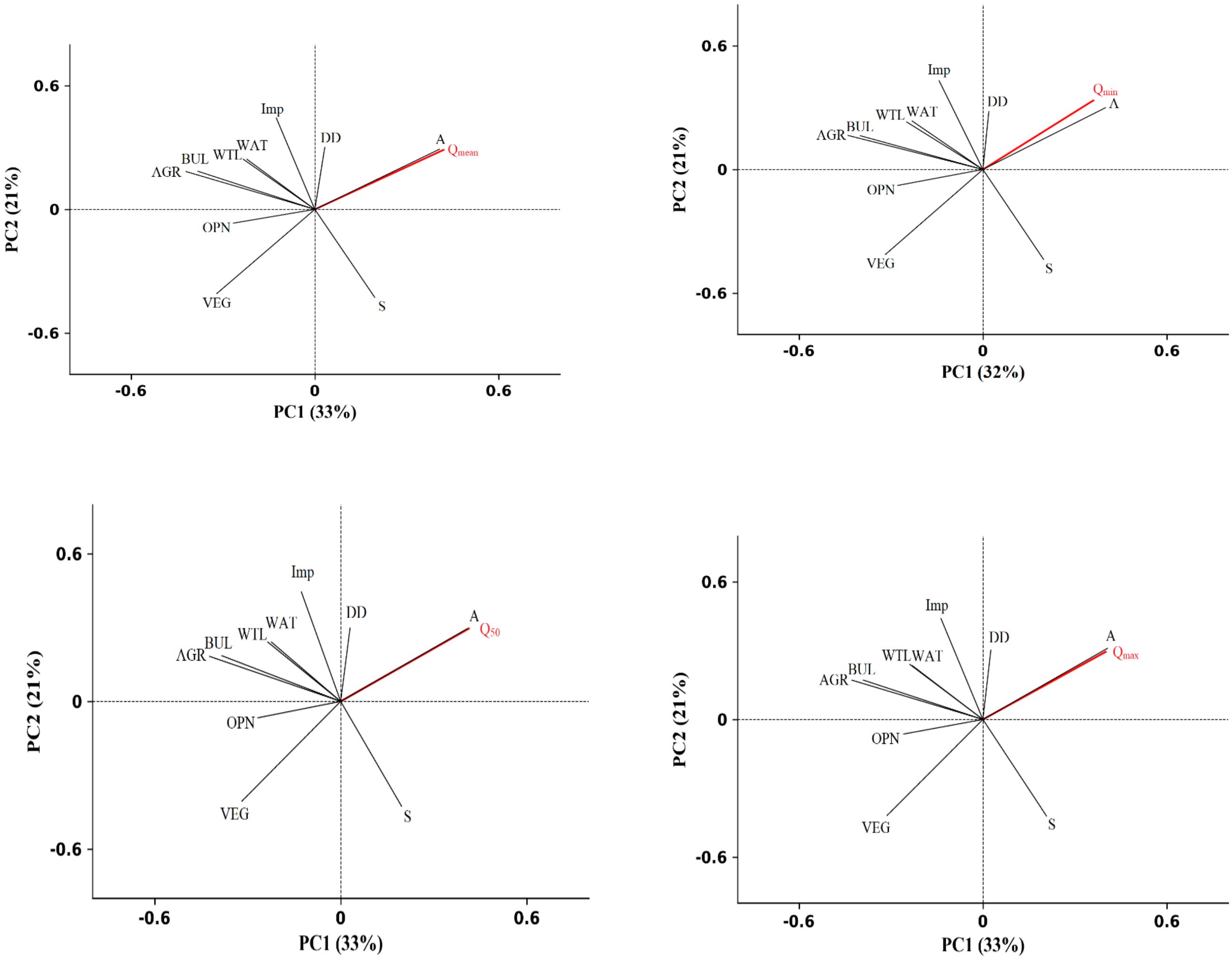

Figure S1. *Biplots from principal component (PC) analysis, showing the interrelations and grouping patterns among climatological variables. Percent variance explained by each component shown in parenthesis. Q, and P refers to different regimes of streamflow and precipitation respectively.*

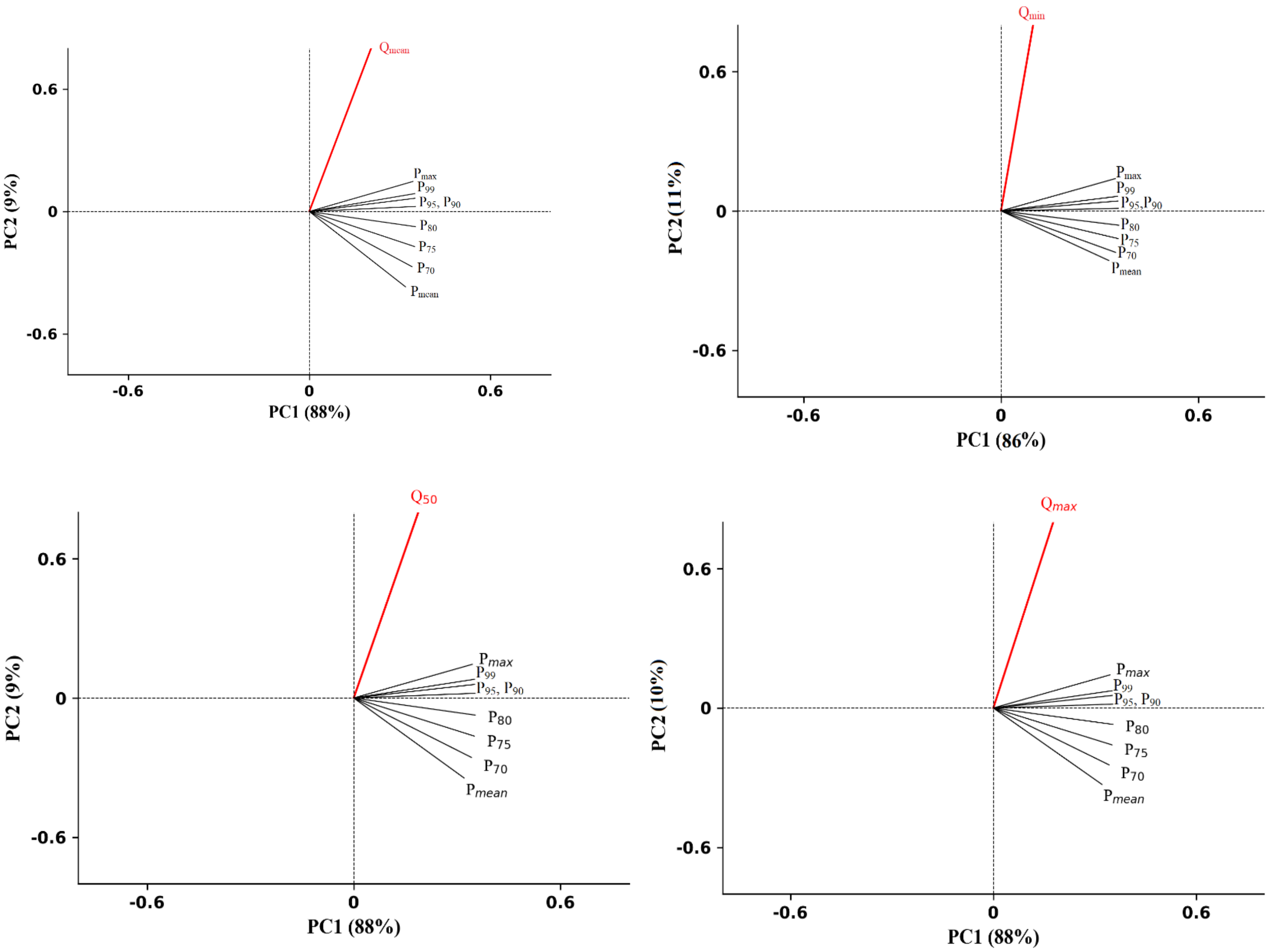

Figure S2. *Biplots from principal component (PC) analysis, showing the interrelations among land cover variables. Percent variance explained by each component shown in parenthesis. WAT, BUL, OPN, VEG, AGR, WTL, Imp, A, S, DD respectively refer to water bodies, built up lands, open lands, vegetated lands, agricultural lands, wetlands, imperviousness, area, mean slope, and drainage density.*

Table S1. Mean of loadings and variance (in parenthesis) of all hydrological, and land cover variables on each principal component (PC) across all study sites.

Varia-bles	PC1 (33%)	PC2 (21%)	PC3 (16%)	PC4 (9%)	PC5 (8%)	PC6 (5%)	PC7 (3%)	PC8 (3%)	PC9 (1%)	PC10 (0%)	PC11 (0%)
Q_1	0.40	0.32	0.03	0.31	0.03	0.26	-0.14	-0.09	0.33	-0.65	-0.08
WAT	-0.22	0.25	-0.44	-0.35	0.07	0.26	-0.60	0.34	-0.12	-0.07	0.00
BUL	-0.39	0.18	0.42	-0.03	0.22	0.13	0.08	0.13	0.02	-0.01	-0.74
OPN	-0.28	-0.06	-0.01	0.63	-0.52	-0.17	-0.15	0.45	-0.05	-0.04	-0.02
VEG	-0.33	-0.40	0.13	0.10	0.12	0.46	-0.15	-0.05	0.60	0.15	0.26
AGR	-0.43	0.19	0.11	0.18	-0.13	0.22	-0.13	-0.64	-0.44	-0.13	0.18
WTL	-0.24	0.25	-0.46	-0.03	-0.16	0.34	0.71	0.10	0.10	-0.06	0.04
Imp	-0.12	0.43	0.46	-0.03	0.27	-0.08	0.13	0.38	-0.05	-0.05	0.58
A	0.41	0.27	0.08	0.27	-0.04	0.48	-0.07	0.06	-0.17	0.64	-0.06
S	0.17	-0.45	0.29	-0.23	-0.22	0.45	0.12	0.27	-0.42	-0.33	0.04
DD	0.03	0.27	0.29	-0.45	-0.71	-0.04	-0.07	-0.11	0.33	0.09	0.00

Note. Q_1, WAT, BUL, OPN, VEG, AGR, WTL, Imp, A, S, DD respectively refer to at-site 1st percentile streamflow, water bodies, built up lands, open lands, vegetated lands, agricultural lands, wetlands, imperviousness, area, mean slope, and drainage density.

Table S2. Mean of loadings and variance (in parenthesis) of all hydrological, and land cover variables on each principal component (PC) across all study sites.

Varia-bles	PC1 (33%)	PC2 (21%)	PC3 (16%)	PC4 (9%)	PC5 (8%)	PC6 (5%)	PC7 (3%)	PC8 (3%)	PC9 (1%)	PC10 (0%)	PC11 (0%)
Q_2	0.41	0.32	0.03	0.31	0.03	0.25	-0.14	-0.09	0.30	-0.67	-0.10
WAT	-0.22	0.25	-0.44	-0.36	0.06	0.25	-0.61	0.33	-0.12	-0.07	0.00
BUL	-0.39	0.18	0.42	-0.03	0.22	0.13	0.08	0.13	0.02	0.00	-0.74
OPN	-0.28	-0.06	-0.01	0.64	-0.51	-0.16	-0.15	0.45	-0.05	-0.04	-0.02
VEG	-0.34	-0.40	0.13	0.09	0.12	0.46	-0.15	-0.05	0.61	0.11	0.26
AGR	-0.43	0.19	0.11	0.18	-0.13	0.21	-0.13	-0.65	-0.44	-0.12	0.18
WTL	-0.23	0.25	-0.46	-0.03	-0.16	0.34	0.71	0.10	0.09	-0.06	0.04
Imp	-0.12	0.43	0.46	-0.03	0.27	-0.08	0.12	0.38	-0.05	-0.06	0.58
A	0.41	0.27	0.08	0.27	-0.04	0.49	-0.07	0.05	-0.13	0.64	-0.04
S	0.17	-0.46	0.29	-0.23	-0.22	0.45	0.12	0.27	-0.43	-0.31	0.03
DD	0.03	0.27	0.29	-0.45	-0.71	-0.04	-0.07	-0.11	0.33	0.08	0.00

Note. Q_2, WAT, BUL, OPN, VEG, AGR, WTL, Imp, A, S, DD respectively refer to at-site 2nd percentile streamflow, water bodies, built up lands, open lands, vegetated lands, agricultural lands, wetlands, imperviousness, area, mean slope, and drainage density.

Table S3. Mean of loadings and variance (in parenthesis) of all hydrological, and land cover variables on each principal component (PC) across all study sites.

Varia-bles	PC1 (33%)	PC2 (21%)	PC3 (16%)	PC4 (9%)	PC5 (8%)	PC6 (5%)	PC7 (3%)	PC8 (3%)	PC9 (1%)	PC10 (0%)	PC11 (0%)
Q_3	0.41	0.32	0.03	0.31	0.03	0.24	-0.13	-0.09	0.29	-0.68	-0.10
WAT	-0.22	0.25	-0.44	-0.36	0.06	0.25	-0.61	0.33	-0.12	-0.07	0.00
BUL	-0.39	0.18	0.42	-0.03	0.22	0.13	0.08	0.13	0.02	0.01	-0.74
OPN	-0.27	-0.06	-0.01	0.64	-0.51	-0.16	-0.15	0.45	-0.05	-0.03	-0.02
VEG	-0.34	-0.40	0.13	0.09	0.12	0.46	-0.15	-0.05	0.62	0.09	0.26
AGR	-0.43	0.19	0.11	0.18	-0.13	0.21	-0.13	-0.65	-0.43	-0.11	0.18
WTL	-0.23	0.25	-0.46	-0.02	-0.16	0.34	0.71	0.11	0.09	-0.06	0.04
Imp	-0.12	0.43	0.47	-0.03	0.28	-0.08	0.12	0.38	-0.05	-0.06	0.58
A	0.41	0.27	0.08	0.27	-0.04	0.49	-0.08	0.04	-0.11	0.64	-0.04
S	0.17	-0.46	0.29	-0.23	-0.22	0.46	0.11	0.27	-0.44	-0.30	0.03
DD	0.03	0.27	0.29	-0.45	-0.71	-0.05	-0.07	-0.11	0.33	0.07	0.00

Note. Q_3, WAT, BUL, OPN, VEG, AGR, WTL, Imp, A, S, DD respectively refer to at-site 3rd percentile streamflow, water bodies, built up lands, open lands, vegetated lands, agricultural lands, wetlands, imperviousness, area, mean slope, and drainage density.

Table S4. Mean of loadings and variance (in parenthesis) of all hydrological, and land cover variables on each principal component (PC) across all study sites.

Varia-bles	PC1 (33%)	PC2 (21%)	PC3 (16%)	PC4 (9%)	PC5 (8%)	PC6 (5%)	PC7 (3%)	PC8 (3%)	PC9 (1%)	PC10 (0%)	PC11 (0%)
Q_4	0.41	0.32	0.03	0.30	0.04	0.24	-0.13	-0.08	0.27	-0.68	-0.10
WAT	-0.22	0.26	-0.44	-0.36	0.06	0.25	-0.61	0.32	-0.12	-0.07	0.00
BUL	-0.39	0.18	0.42	-0.03	0.22	0.13	0.08	0.13	0.02	0.00	-0.74
OPN	-0.27	-0.06	-0.01	0.64	-0.50	-0.16	-0.16	0.45	-0.05	-0.03	-0.02
VEG	-0.34	-0.40	0.13	0.09	0.12	0.46	-0.15	-0.05	0.62	0.07	0.25
AGR	-0.43	0.19	0.11	0.18	-0.13	0.21	-0.13	-0.65	-0.43	-0.11	0.18
WTL	-0.23	0.25	-0.46	-0.02	-0.16	0.35	0.71	0.11	0.08	-0.06	0.04
Imp	-0.12	0.43	0.47	-0.03	0.28	-0.07	0.12	0.38	-0.05	-0.06	0.58
A	0.41	0.27	0.08	0.27	-0.03	0.50	-0.09	0.04	-0.10	0.64	-0.04
S	0.17	-0.46	0.29	-0.23	-0.23	0.46	0.11	0.27	-0.45	-0.29	0.03
DD	0.03	0.27	0.29	-0.44	-0.72	-0.05	-0.06	-0.10	0.33	0.07	0.00

Note. Q_4, WAT, BUL, OPN, VEG, AGR, WTL, Imp, A, S, DD respectively refer to at-site 4th percentile streamflow, water bodies, built up lands, open lands, vegetated lands, agricultural lands, wetlands, imperviousness, area, mean slope, and drainage density.

Table S5. Mean of loadings and variance (in parenthesis) of all hydrological, and land cover variables on each principal component (PC) across all study sites.

Varia-bles	PC1 (33%)	PC2 (21%)	PC3 (16%)	PC4 (9%)	PC5 (8%)	PC6 (5%)	PC7 (3%)	PC8 (3%)	PC9 (1%)	PC10 (0%)	PC11 (0%)
Q_5	0.41	0.32	0.03	0.30	0.04	0.24	-0.13	-0.08	0.26	-0.69	-0.11
WAT	-0.22	0.26	-0.44	-0.36	0.05	0.25	-0.62	0.32	-0.12	-0.07	0.00
BUL	-0.39	0.18	0.42	-0.04	0.22	0.13	0.08	0.14	0.02	0.00	-0.74
OPN	-0.27	-0.06	-0.01	0.65	-0.50	-0.16	-0.16	0.45	-0.06	-0.03	-0.02
VEG	-0.34	-0.40	0.13	0.09	0.11	0.45	-0.15	-0.05	0.63	0.06	0.25
AGR	-0.42	0.19	0.11	0.18	-0.13	0.21	-0.12	-0.66	-0.43	-0.11	0.18
WTL	-0.23	0.25	-0.46	-0.02	-0.16	0.35	0.70	0.12	0.08	-0.06	0.04
Imp	-0.12	0.43	0.47	-0.03	0.28	-0.07	0.11	0.38	-0.05	-0.06	0.58
A	0.41	0.27	0.08	0.27	-0.03	0.50	-0.09	0.03	-0.09	0.64	-0.04
S	0.17	-0.46	0.29	-0.23	-0.23	0.46	0.11	0.27	-0.45	-0.28	0.03
DD	0.03	0.27	0.29	-0.43	-0.72	-0.05	-0.06	-0.10	0.33	0.06	0.00

Note. Q_5, WAT, BUL, OPN, VEG, AGR, WTL, Imp, A, S, DD respectively refer to at-site 5th percentile streamflow, water bodies, built up lands, open lands, vegetated lands, agricultural lands, wetlands, imperviousness, area, mean slope, and drainage density.

Table S6. Mean of loadings and variance (in parenthesis) of all hydrological, and land cover variables on each principal component (PC) across all study sites.

Varia-bles	PC1 (33%)	PC2 (21%)	PC3 (16%)	PC4 (9%)	PC5 (8%)	PC6 (5%)	PC7 (3%)	PC8 (3%)	PC9 (1%)	PC10 (0%)	PC11 (0%)
Q_{10}	0.42	0.31	0.03	0.29	0.04	0.25	-0.11	-0.07	-0.22	-0.71	-0.13
WAT	-0.21	0.26	-0.44	-0.36	0.04	0.25	-0.63	0.30	0.13	-0.06	0.00
BUL	-0.39	0.18	0.42	-0.04	0.22	0.14	0.07	0.14	-0.01	0.03	-0.74
OPN	-0.27	-0.05	-0.01	0.67	-0.47	-0.16	-0.17	0.44	0.06	-0.02	-0.02
VEG	-0.34	-0.40	0.13	0.08	0.11	0.46	-0.15	-0.05	-0.63	0.01	0.25
AGR	-0.42	0.19	0.11	0.19	-0.12	0.21	-0.11	-0.67	0.43	-0.10	0.18
WTL	-0.23	0.26	-0.46	-0.01	-0.16	0.35	0.70	0.14	-0.07	-0.05	0.04
Imp	-0.12	0.43	0.47	-0.04	0.28	-0.07	0.10	0.38	0.06	-0.07	0.58
A	0.41	0.27	0.08	0.28	-0.02	0.49	-0.10	0.01	0.04	0.65	-0.01
S	0.17	-0.46	0.29	-0.22	-0.24	0.46	0.10	0.27	0.48	-0.23	0.03
DD	0.03	0.28	0.29	-0.40	-0.74	-0.06	-0.05	-0.10	-0.33	0.04	0.00

Note. Q_{10}, WAT, BUL, OPN, VEG, AGR, WTL, Imp, A, S, DD respectively refer to at-site 10th percentile streamflow, water bodies, built up lands, open lands, vegetated lands, agricultural lands, wetlands, imperviousness, area, mean slope, and drainage density.

Table S7. Mean of loadings and variance (in parenthesis) of all hydrological, and land cover variables on each principal component (PC) across all study sites.

Varia-bles	PC1 (33%)	PC2 (21%)	PC3 (16%)	PC4 (9%)	PC5 (8%)	PC6 (5%)	PC7 (3%)	PC8 (3%)	PC9 (1%)	PC10 (0%)	PC11 (0%)
Q_{20}	0.42	0.30	0.04	0.30	0.05	0.27	-0.10	-0.05	-0.17	-0.07	-0.72
WAT	-0.22	0.25	-0.44	-0.35	0.01	0.28	-0.63	0.27	0.13	0.00	-0.04
BUL	-0.39	0.18	0.42	-0.05	0.21	0.15	0.07	0.14	0.00	-0.74	-0.02
OPN	-0.27	-0.06	-0.01	0.70	-0.42	-0.18	-0.19	0.43	0.07	-0.02	0.00
VEG	-0.33	-0.40	0.14	0.08	0.11	0.46	-0.13	-0.05	-0.63	0.25	0.00
AGR	-0.42	0.19	0.11	0.20	-0.11	0.20	-0.08	-0.68	0.42	0.19	-0.07
WTL	-0.23	0.25	-0.46	0.02	-0.16	0.35	0.70	0.16	-0.05	0.05	-0.02
Imp	-0.12	0.43	0.46	-0.06	0.28	-0.07	0.08	0.38	0.06	0.58	-0.01
A	0.41	0.28	0.08	0.29	0.01	0.44	-0.11	-0.03	0.00	-0.07	0.67
S	0.18	-0.45	0.29	-0.19	-0.26	0.47	0.10	0.27	0.51	0.05	-0.14
DD	0.03	0.28	0.29	-0.36	-0.76	-0.05	-0.04	-0.09	-0.34	-0.01	0.02

Note. Q_{20}, WAT, BUL, OPN, VEG, AGR, WTL, Imp, A, S, DD respectively refer to at-site 20th percentile streamflow, water bodies, built up lands, open lands, vegetated lands, agricultural lands, wetlands, imperviousness, area, mean slope, and drainage density.

Table S8. Mean of loadings and variance (in parenthesis) of all hydrological, and land cover variables on each principal component (PC) across all study sites.

Varia-bles	PC1 (33%)	PC2 (21%)	PC3 (16%)	PC4 (9%)	PC5 (8%)	PC6 (5%)	PC7 (3%)	PC8 (3%)	PC9 (1%)	PC10 (0%)	PC11 (0%)
Q_{30}	0.42	0.30	0.04	0.31	0.05	0.30	-0.09	-0.04	-0.13	-0.13	-0.71
WAT	-0.22	0.25	-0.44	-0.33	-0.01	0.30	-0.64	0.26	0.13	0.00	-0.03
BUL	-0.39	0.18	0.42	-0.06	0.21	0.16	0.07	0.15	0.00	-0.74	0.05
OPN	-0.27	-0.06	0.00	0.71	-0.39	-0.20	-0.20	0.43	0.07	-0.02	0.01
VEG	-0.33	-0.40	0.15	0.09	0.10	0.47	-0.12	-0.04	-0.62	0.25	-0.02
AGR	-0.42	0.19	0.12	0.21	-0.11	0.19	-0.07	-0.68	0.42	0.18	-0.07
WTL	-0.23	0.25	-0.46	0.04	-0.16	0.34	0.70	0.18	-0.05	0.05	0.00
Imp	-0.12	0.44	0.46	-0.08	0.28	-0.07	0.07	0.38	0.06	0.58	-0.05
A	0.41	0.29	0.07	0.30	0.03	0.40	-0.11	-0.05	-0.03	-0.01	0.69
S	0.19	-0.44	0.29	-0.16	-0.27	0.47	0.10	0.27	0.52	0.05	-0.08
DD	0.03	0.29	0.28	-0.33	-0.78	-0.05	-0.04	-0.09	-0.34	-0.01	0.01

Note. Q_{30}, WAT, BUL, OPN, VEG, AGR, WTL, Imp, A, S, DD respectively refer to at-site 30th percentile streamflow, water bodies, built up lands, open lands, vegetated lands, agricultural lands, wetlands, imperviousness, area, mean slope, and drainage density.

Table S9. Mean of loadings and variance (in parenthesis) of all hydrological, and land cover variables on each principal component (PC) across all study sites.

Varia-bles	PC1 (33%)	PC2 (21%)	PC3 (16%)	PC4 (9%)	PC5 (8%)	PC6 (5%)	PC7 (3%)	PC8 (3%)	PC9 (1%)	PC10 (0%)	PC11 (0%)
Q_{40}	0.41	0.30	0.05	0.32	0.05	0.32	-0.08	-0.03	-0.12	-0.14	-0.70
WAT	-0.22	0.24	-0.44	-0.32	-0.02	0.32	-0.64	0.26	0.14	0.00	-0.02
BUL	-0.38	0.18	0.42	-0.06	0.20	0.16	0.07	0.15	0.00	-0.74	0.07
OPN	-0.27	-0.06	0.00	0.71	-0.37	-0.22	-0.21	0.42	0.07	-0.01	0.01
VEG	-0.32	-0.41	0.16	0.10	0.10	0.47	-0.12	-0.04	-0.62	0.25	-0.02
AGR	-0.42	0.18	0.12	0.21	-0.10	0.18	-0.06	-0.68	0.42	0.18	-0.06
WTL	-0.23	0.25	-0.46	0.06	-0.15	0.33	0.70	0.18	-0.04	0.05	0.01
Imp	-0.13	0.44	0.45	-0.09	0.28	-0.07	0.07	0.38	0.06	0.58	-0.06
A	0.41	0.29	0.07	0.31	0.04	0.37	-0.11	-0.06	-0.05	0.01	0.70
S	0.19	-0.43	0.30	-0.15	-0.28	0.47	0.10	0.27	0.53	0.06	-0.05
DD	0.03	0.29	0.27	-0.31	-0.79	-0.04	-0.04	-0.09	-0.33	-0.01	0.01

Note. Q_{40}, WAT, BUL, OPN, VEG, AGR, WTL, Imp, A, S, DD respectively refer to at-site 40th percentile streamflow, water bodies, built up lands, open lands, vegetated lands, agricultural lands, wetlands, imperviousness, area, mean slope, and drainage density.

Table S10. Mean of loadings and variance (in parenthesis) of all hydrological, and land cover variables on each principal component (PC) across all study sites.

Varia-bles	PC1 (33%)	PC2 (21%)	PC3 (16%)	PC4 (9%)	PC5 (8%)	PC6 (5%)	PC7 (3%)	PC8 (3%)	PC9 (1%)	PC10 (0%)	PC11 (0%)
Q_{60}	0.42	0.29	0.05	0.32	0.06	0.32	-0.08	-0.03	-0.09	-0.15	-0.70
WAT	-0.22	0.24	-0.44	-0.31	-0.04	0.32	-0.64	0.26	0.13	0.00	-0.02
BUL	-0.38	0.19	0.42	-0.07	0.20	0.16	0.07	0.15	0.01	-0.73	0.08
OPN	-0.27	-0.07	0.01	0.73	-0.34	-0.23	-0.21	0.42	0.07	-0.01	0.01
VEG	-0.32	-0.41	0.17	0.09	0.10	0.47	-0.11	-0.04	-0.62	0.25	-0.04
AGR	-0.42	0.18	0.12	0.22	-0.09	0.18	-0.07	-0.68	0.42	0.19	-0.04
WTL	-0.24	0.24	-0.46	0.07	-0.15	0.33	0.70	0.18	-0.04	0.05	0.01
Imp	-0.13	0.45	0.45	-0.10	0.28	-0.07	0.07	0.38	0.06	0.58	-0.07
A	0.41	0.30	0.07	0.31	0.06	0.36	-0.11	-0.05	-0.08	0.02	0.70
S	0.20	-0.43	0.30	-0.14	-0.29	0.46	0.10	0.27	0.53	0.07	-0.02
DD	0.03	0.30	0.27	-0.28	-0.80	-0.05	-0.03	-0.08	-0.33	-0.01	0.00

Note. Q_{60}, WAT, BUL, OPN, VEG, AGR, WTL, Imp, A, S, DD respectively refer to at-site 60th percentile streamflow, water bodies, built up lands, open lands, vegetated lands, agricultural lands, wetlands, imperviousness, area, mean slope, and drainage density.

Table S11. Mean of loadings and variance (in parenthesis) of all hydrological, and land cover variables on each principal component (PC) across all study sites.

Varia-bles	PC1 (33%)	PC2 (21%)	PC3 (16%)	PC4 (9%)	PC5 (8%)	PC6 (5%)	PC7 (3%)	PC8 (3%)	PC9 (1%)	PC10 (0%)	PC11 (0%)
Q_{70}	0.42	0.29	0.05	0.30	0.06	0.32	-0.08	-0.03	-0.06	-0.25	0.68
WAT	-0.22	0.24	-0.44	-0.32	-0.06	0.31	-0.64	0.26	0.13	0.00	0.02
BUL	-0.38	0.19	0.42	-0.08	0.19	0.16	0.07	0.15	0.01	-0.72	-0.18
OPN	-0.27	-0.07	0.01	0.75	-0.30	-0.22	-0.21	0.42	0.07	-0.01	-0.01
VEG	-0.32	-0.41	0.16	0.07	0.09	0.47	-0.12	-0.04	-0.63	0.22	0.12
AGR	-0.42	0.18	0.12	0.22	-0.09	0.19	-0.07	-0.69	0.41	0.18	0.06
WTL	-0.23	0.24	-0.46	0.08	-0.15	0.34	0.70	0.18	-0.04	0.05	0.00
Imp	-0.13	0.44	0.45	-0.11	0.28	-0.07	0.07	0.38	0.06	0.56	0.15
A	0.41	0.29	0.07	0.31	0.08	0.37	-0.12	-0.06	-0.12	0.12	-0.68
S	0.20	-0.43	0.30	-0.13	-0.30	0.46	0.09	0.27	0.53	0.07	0.00
DD	0.03	0.30	0.27	-0.24	-0.81	-0.07	-0.03	-0.08	-0.33	-0.01	0.00

Note. Q_{70}, WAT, BUL, OPN, VEG, AGR, WTL, Imp, A, S, DD respectively refer to at-site 70th percentile streamflow, water bodies, built up lands, open lands, vegetated lands, agricultural lands, wetlands, imperviousness, area, mean slope, and drainage density.

Table S12. Mean of loadings and variance (in parenthesis) of all hydrological, and land cover variables on each principal component (PC) across all study sites.

Varia-bles	PC1 (33%)	PC2 (21%)	PC3 (16%)	PC4 (9%)	PC5 (8%)	PC6 (5%)	PC7 (3%)	PC8 (3%)	PC9 (1%)	PC10 (0%)	PC11 (0%)
Q_{80}	0.42	0.29	0.05	0.30	0.06	0.32	-0.08	-0.03	-0.06	-0.23	0.69
WAT	-0.22	0.24	-0.44	-0.32	-0.06	0.31	-0.64	0.26	0.13	0.00	0.02
BUL	-0.38	0.19	0.42	-0.08	0.19	0.16	0.07	0.15	0.01	-0.72	-0.16
OPN	-0.27	-0.07	0.01	0.75	-0.30	-0.22	-0.21	0.42	0.07	-0.01	-0.01
VEG	-0.32	-0.41	0.16	0.07	0.09	0.47	-0.12	-0.04	-0.63	0.23	0.11
AGR	-0.42	0.18	0.12	0.22	-0.09	0.19	-0.07	-0.69	0.41	0.19	0.05
WTL	-0.23	0.24	-0.46	0.08	-0.15	0.34	0.70	0.18	-0.04	0.05	0.00
Imp	-0.13	0.45	0.45	-0.11	0.28	-0.07	0.07	0.38	0.06	0.57	0.13
A	0.41	0.29	0.07	0.31	0.08	0.37	-0.12	-0.05	-0.12	0.09	-0.68
S	0.20	-0.43	0.30	-0.13	-0.30	0.46	0.10	0.27	0.53	0.07	0.00
DD	0.03	0.30	0.27	-0.24	-0.81	-0.07	-0.03	-0.08	-0.33	-0.01	0.00

Note. Q_{80}, WAT, BUL, OPN, VEG, AGR, WTL, Imp, A, S, DD respectively refer to at-site 80th percentile streamflow, water bodies, built up lands, open lands, vegetated lands, agricultural lands, wetlands, imperviousness, area, mean slope, and drainage density.

Table S13. Mean of loadings and variance (in parenthesis) of all hydrological, and land cover variables on each principal component (PC) across all study sites.

Varia-bles	PC1 (33%)	PC2 (21%)	PC3 (16%)	PC4 (9%)	PC5 (8%)	PC6 (5%)	PC7 (3%)	PC8 (3%)	PC9 (1%)	PC10 (0%)	PC11 (0%)
Q_{90}	0.42	0.29	0.05	0.30	0.06	0.32	-0.08	-0.03	-0.06	-0.23	0.69
WAT	-0.22	0.24	-0.44	-0.32	-0.06	0.31	-0.64	0.26	0.13	0.00	0.02
BUL	-0.38	0.19	0.42	-0.08	0.19	0.17	0.07	0.15	0.01	-0.72	-0.16
OPN	-0.27	-0.07	0.01	0.75	-0.29	-0.22	-0.21	0.42	0.07	-0.01	-0.01
VEG	-0.32	-0.41	0.16	0.07	0.09	0.47	-0.12	-0.04	-0.63	0.23	0.11
AGR	-0.42	0.18	0.12	0.22	-0.08	0.19	-0.07	-0.69	0.41	0.19	0.05
WTL	-0.23	0.24	-0.46	0.08	-0.15	0.34	0.70	0.18	-0.04	0.05	0.00
Imp	-0.13	0.45	0.45	-0.11	0.28	-0.07	0.07	0.38	0.06	0.57	0.13
A	0.41	0.29	0.07	0.31	0.08	0.37	-0.12	-0.06	-0.12	0.10	-0.68
S	0.20	-0.43	0.30	-0.13	-0.30	0.46	0.09	0.27	0.53	0.07	0.00
DD	0.03	0.30	0.27	-0.23	-0.81	-0.07	-0.03	-0.08	-0.33	-0.01	0.00

Note. Q_{90}, WAT, BUL, OPN, VEG, AGR, WTL, Imp, A, S, DD respectively refer to at-site 90th percentile streamflow, water bodies, built up lands, open lands, vegetated lands, agricultural lands, wetlands, imperviousness, area, mean slope, and drainage density.

Table S14. Mean of loadings and variance (in parenthesis) of all hydrological, and land cover variables on each principal component (PC) across all study sites.

Varia-bles	PC1 (33%)	PC2 (21%)	PC3 (16%)	PC4 (9%)	PC5 (8%)	PC6 (5%)	PC7 (3%)	PC8 (3%)	PC9 (1%)	PC10 (0%)	PC11 (0%)
Q_{95}	0.42	0.28	0.05	0.30	0.06	0.32	-0.08	-0.03	-0.04	-0.23	0.69
WAT	-0.22	0.24	-0.44	-0.31	-0.08	0.31	-0.64	0.25	0.13	0.00	0.02
BUL	-0.38	0.18	0.42	-0.09	0.19	0.17	0.07	0.15	0.01	-0.72	-0.16
OPN	-0.27	-0.07	0.01	0.76	-0.26	-0.22	-0.21	0.42	0.07	-0.01	-0.02
VEG	-0.32	-0.41	0.16	0.07	0.08	0.46	-0.12	-0.04	-0.63	0.22	0.12
AGR	-0.42	0.18	0.12	0.23	-0.08	0.18	-0.06	-0.69	0.41	0.19	0.05
WTL	-0.23	0.24	-0.46	0.09	-0.15	0.34	0.70	0.18	-0.04	0.05	-0.01
Imp	-0.13	0.44	0.45	-0.12	0.28	-0.06	0.06	0.38	0.06	0.57	0.13
A	0.41	0.29	0.07	0.31	0.10	0.37	-0.13	-0.07	-0.14	0.09	-0.68
S	0.20	-0.43	0.30	-0.12	-0.31	0.46	0.09	0.27	0.53	0.08	-0.02
DD	0.03	0.31	0.27	-0.20	-0.82	-0.08	-0.03	-0.08	-0.33	-0.01	0.00

Note. Q_{95}, WAT, BUL, OPN, VEG, AGR, WTL, Imp, A, S, DD respectively refer to at-site 95th percentile streamflow, water bodies, built up lands, open lands, vegetated lands, agricultural lands, wetlands, imperviousness, area, mean slope, and drainage density.

Table S15. Mean of loadings and variance (in parenthesis) of all hydrological, and land cover variables on each principal component (PC) across all study sites.

Varia-bles	PC1 (33%)	PC2 (21%)	PC3 (16%)	PC4 (9%)	PC5 (8%)	PC6 (5%)	PC7 (3%)	PC8 (3%)	PC9 (1%)	PC10 (0%)	PC11 (0%)
Q_{96}	0.42	0.28	0.05	0.30	0.06	0.32	-0.08	-0.03	-0.04	-0.23	0.70
WAT	-0.22	0.24	-0.44	-0.31	-0.08	0.31	-0.64	0.25	0.13	0.00	0.01
BUL	-0.38	0.18	0.42	-0.09	0.18	0.17	0.07	0.15	0.01	-0.72	-0.15
OPN	-0.26	-0.07	0.01	0.77	-0.25	-0.22	-0.21	0.42	0.07	-0.01	-0.02
VEG	-0.32	-0.41	0.16	0.06	0.08	0.46	-0.12	-0.04	-0.63	0.22	0.13
AGR	-0.42	0.18	0.12	0.23	-0.08	0.18	-0.06	-0.69	0.41	0.19	0.05
WTL	-0.23	0.24	-0.46	0.09	-0.15	0.34	0.70	0.19	-0.04	0.05	-0.01
Imp	-0.13	0.44	0.45	-0.12	0.28	-0.06	0.06	0.38	0.06	0.57	0.13
A	0.41	0.29	0.07	0.30	0.10	0.37	-0.12	-0.07	-0.14	0.09	-0.68
S	0.20	-0.43	0.30	-0.11	-0.31	0.46	0.09	0.27	0.53	0.08	-0.03
DD	0.03	0.31	0.27	-0.19	-0.82	-0.09	-0.02	-0.08	-0.33	-0.01	0.01

Note. Q_{96}, WAT, BUL, OPN, VEG, AGR, WTL, Imp, A, S, DD respectively refer to at-site 96th percentile streamflow, water bodies, built up lands, open lands, vegetated lands, agricultural lands, wetlands, imperviousness, area, mean slope, and drainage density.

Table S16. Mean of loadings and variance (in parenthesis) of all hydrological, and land cover variables on each principal component (PC) across all study sites.

Varia-bles	PC1 (33%)	PC2 (21%)	PC3 (16%)	PC4 (9%)	PC5 (8%)	PC6 (5%)	PC7 (3%)	PC8 (3%)	PC9 (1%)	PC10 (0%)	PC11 (0%)
Q_{97}	0.42	0.28	0.05	0.30	0.06	0.32	-0.08	-0.03	-0.03	-0.22	-0.70
WAT	-0.22	0.24	-0.44	-0.31	-0.09	0.32	-0.64	0.25	0.13	0.01	-0.01
BUL	-0.38	0.18	0.42	-0.09	0.18	0.17	0.07	0.15	0.01	-0.73	0.14
OPN	-0.26	-0.07	0.01	0.77	-0.25	-0.22	-0.21	0.42	0.07	-0.01	0.02
VEG	-0.32	-0.41	0.16	0.06	0.08	0.46	-0.11	-0.04	-0.63	0.22	-0.13
AGR	-0.42	0.18	0.12	0.23	-0.08	0.18	-0.06	-0.69	0.41	0.19	-0.04
WTL	-0.23	0.24	-0.46	0.09	-0.15	0.34	0.70	0.19	-0.04	0.05	0.01
Imp	-0.13	0.44	0.45	-0.12	0.28	-0.06	0.06	0.38	0.06	0.57	-0.12
A	0.41	0.29	0.07	0.30	0.10	0.37	-0.12	-0.07	-0.15	0.08	0.68
S	0.20	-0.42	0.30	-0.11	-0.32	0.46	0.09	0.27	0.53	0.08	0.03
DD	0.03	0.31	0.27	-0.19	-0.82	-0.09	-0.02	-0.08	-0.33	-0.01	-0.01

Note. Q_{97}, WAT, BUL, OPN, VEG, AGR, WTL, Imp, A, S, DD respectively refer to at-site 97th percentile streamflow, water bodies, built up lands, open lands, vegetated lands, agricultural lands, wetlands, imperviousness, area, mean slope, and drainage density.

Table S17. Mean of loadings and variance (in parenthesis) of all hydrological, and land cover variables on each principal component (PC) across all study sites.

Varia-bles	PC1 (33%)	PC2 (21%)	PC3 (16%)	PC4 (9%)	PC5 (8%)	PC6 (5%)	PC7 (3%)	PC8 (3%)	PC9 (1%)	PC10 (0%)	PC11 (0%)
Q_{98}	0.42	0.28	0.05	0.30	0.06	0.32	-0.08	-0.02	-0.03	-0.20	-0.70
WAT	-0.22	0.24	-0.44	-0.31	-0.09	0.32	-0.64	0.25	0.13	0.01	-0.01
BUL	-0.38	0.18	0.42	-0.09	0.18	0.17	0.07	0.15	0.01	-0.73	0.13
OPN	-0.26	-0.07	0.01	0.77	-0.24	-0.22	-0.21	0.42	0.07	-0.01	0.03
VEG	-0.32	-0.41	0.16	0.06	0.08	0.46	-0.11	-0.05	-0.63	0.23	-0.12
AGR	-0.42	0.18	0.12	0.23	-0.08	0.18	-0.06	-0.69	0.41	0.19	-0.04
WTL	-0.23	0.24	-0.46	0.09	-0.15	0.34	0.70	0.19	-0.04	0.05	0.01
Imp	-0.13	0.44	0.45	-0.12	0.27	-0.06	0.06	0.38	0.06	0.57	-0.11
A	0.41	0.30	0.06	0.30	0.11	0.36	-0.12	-0.08	-0.15	0.07	0.68
S	0.20	-0.42	0.30	-0.11	-0.32	0.46	0.09	0.27	0.53	0.08	0.04
DD	0.03	0.31	0.27	-0.18	-0.82	-0.09	-0.02	-0.08	-0.33	-0.01	-0.01

Note. Q_{98}, WAT, BUL, OPN, VEG, AGR, WTL, Imp, A, S, DD respectively refer to at-site 98th percentile streamflow, water bodies, built up lands, open lands, vegetated lands, agricultural lands, wetlands, imperviousness, area, mean slope, and drainage density.

Table S18. Mean of loadings and variance (in parenthesis) of all hydrological, and land cover variables on each principal component (PC) across all study sites.

Varia-bles	PC1 (33%)	PC2 (21%)	PC3 (16%)	PC4 (9%)	PC5 (8%)	PC6 (5%)	PC7 (3%)	PC8 (3%)	PC9 (1%)	PC10 (0%)	PC11 (0%)
Q_{99}	0.42	0.28	0.05	0.30	0.06	0.32	-0.08	-0.02	-0.02	-0.23	-0.69
WAT	-0.22	0.24	-0.44	-0.30	-0.10	0.32	-0.64	0.24	0.13	0.01	-0.01
BUL	-0.38	0.18	0.42	-0.09	0.18	0.17	0.06	0.15	0.02	-0.72	0.16
OPN	-0.26	-0.07	0.01	0.77	-0.23	-0.22	-0.21	0.42	0.07	-0.01	0.03
VEG	-0.32	-0.42	0.16	0.06	0.07	0.46	-0.11	-0.05	-0.63	0.22	-0.14
AGR	-0.42	0.18	0.12	0.23	-0.07	0.18	-0.06	-0.69	0.41	0.19	-0.05
WTL	-0.23	0.24	-0.46	0.10	-0.15	0.34	0.70	0.19	-0.04	0.05	0.01
Imp	-0.13	0.44	0.45	-0.13	0.27	-0.05	0.06	0.38	0.05	0.57	-0.13
A	0.40	0.30	0.07	0.30	0.11	0.36	-0.12	-0.08	-0.16	0.09	0.67
S	0.20	-0.42	0.30	-0.10	-0.32	0.46	0.09	0.26	0.53	0.09	0.05
DD	0.03	0.31	0.27	-0.17	-0.82	-0.09	-0.02	-0.08	-0.33	-0.01	-0.01

Note. Q_{99}, WAT, BUL, OPN, VEG, AGR, WTL, Imp, A, S, DD respectively refer to at-site 99th percentile streamflow, water bodies, built up lands, open lands, vegetated lands, agricultural lands, wetlands, imperviousness, area, mean slope, and drainage density.

Table S19. Mean of loadings and variance (in parenthesis) of all climatological variables on each principal component (PC) across all study sites.

Variables	PC1 (87%)	PC2 (11%)	PC3 (2%)	PC4 (1%)	PC5 (0%)	PC6 (0%)	PC7 (0%)	PC8 (0%)	PC9 (0%)
Q_1	0.17	0.90	0.35	-0.17	0.05	0.02	0.00	0.01	0.00
P_{mean}	0.32	-0.27	0.70	0.20	-0.49	0.17	-0.13	0.06	0.00
P_{70}	0.35	-0.23	0.29	-0.21	0.47	-0.57	-0.03	-0.19	-0.35
P_{75}	0.35	-0.15	0.08	-0.19	0.27	0.23	0.52	-0.09	0.64
P_{80}	0.36	-0.08	-0.15	-0.21	0.30	0.35	-0.49	0.59	-0.04
P_{90}	0.35	0.01	-0.26	-0.15	-0.19	0.40	0.42	-0.15	-0.63
P_{95}	0.35	0.04	-0.31	-0.12	-0.21	0.01	-0.52	-0.63	0.23
P_{99}	0.35	0.07	-0.30	-0.09	-0.47	-0.56	0.18	0.42	0.15
P_{max}	0.35	0.15	-0.16	0.87	0.26	-0.02	0.03	0.00	-0.01

Note. Q_1, P_{mean}, P_{max} respectively refer to at-site 1st streamflow, at-site mean, and maximum daily precipitation. The rest of the variables represent percentiles of at-site daily precipitation.

Table S20. Mean of loadings and variance (in parenthesis) of all climatological variables on each principal component (PC) across all study sites.

Variables	PC1 (87%)	PC2 (10%)	PC3 (2%)	PC4 (1%)	PC5 (0%)	PC6 (0%)	PC7 (0%)	PC8 (0%)	PC9 (0%)
Q_2	0.18	0.90	0.35	-0.18	0.05	0.02	0.00	0.01	0.00
P_{mean}	0.32	-0.28	0.70	0.19	-0.49	0.16	-0.13	0.06	0.00
P_{70}	0.34	-0.23	0.28	-0.21	0.47	-0.57	-0.03	-0.19	-0.35
P_{75}	0.35	-0.16	0.07	-0.19	0.27	0.23	0.52	-0.09	0.64
P_{80}	0.35	-0.08	-0.15	-0.21	0.30	0.34	-0.49	0.59	-0.03
P_{90}	0.35	0.01	-0.26	-0.14	-0.19	0.40	0.42	-0.15	-0.63
P_{95}	0.35	0.05	-0.31	-0.12	-0.21	0.01	-0.52	-0.63	0.23
P_{99}	0.35	0.07	-0.30	-0.09	-0.47	-0.57	0.19	0.42	0.15
P_{max}	0.35	0.15	-0.15	0.88	0.26	-0.02	0.04	0.00	-0.01

Note. Q_2, P_{mean}, P_{max} respectively refer to at-site 2nd streamflow, at-site mean, and maximum daily precipitation. The rest of the variables represent percentiles of at-site daily precipitation.

Table S21. Mean of loadings and variance (in parenthesis) of all climatological variables on each principal component (PC) across all study sites.

Variables	PC1 (87%)	PC2 (10%)	PC3 (2%)	PC4 (1%)	PC5 (0%)	PC6 (0%)	PC7 (0%)	PC8 (0%)	PC9 (0%)
Q_3	0.18	0.90	0.36	-0.18	0.05	0.02	0.00	0.01	0.00
P_{mean}	0.32	-0.28	0.70	0.19	-0.49	0.16	-0.13	0.06	0.00
P_{70}	0.34	-0.24	0.28	-0.21	0.48	-0.57	-0.03	-0.19	-0.35
P_{75}	0.35	-0.16	0.07	-0.19	0.27	0.23	0.52	-0.09	0.64
P_{80}	0.35	-0.08	-0.15	-0.21	0.30	0.35	-0.49	0.59	-0.03
P_{90}	0.35	0.01	-0.26	-0.14	-0.19	0.40	0.42	-0.15	-0.63
P_{95}	0.35	0.05	-0.31	-0.12	-0.21	0.01	-0.52	-0.63	0.23
P_{99}	0.35	0.07	-0.30	-0.09	-0.47	-0.57	0.18	0.42	0.15
P_{max}	0.35	0.15	-0.15	0.88	0.25	-0.02	0.03	0.00	-0.01

Note. Q_3, P_{mean}, P_{max} respectively refer to at-site 3rd streamflow, at-site mean, and maximum daily precipitation. The rest of the variables represent percentiles of at-site daily precipitation.

Table S22. Mean of loadings and variance (in parenthesis) of all climatological variables on each principal component (PC) across all study sites.

Variables	PC1 (87%)	PC2 (10%)	PC3 (2%)	PC4 (1%)	PC5 (0%)	PC6 (0%)	PC7 (0%)	PC8 (0%)	PC9 (0%)
Q_4	0.18	0.89	0.36	-0.18	0.06	0.02	0.00	0.01	0.00
P_{mean}	0.32	-0.29	0.70	0.19	-0.49	0.16	-0.13	0.06	0.00
P_{70}	0.34	-0.24	0.27	-0.21	0.48	-0.56	-0.03	-0.19	-0.35
P_{75}	0.35	-0.16	0.07	-0.19	0.27	0.23	0.52	-0.09	0.64
P_{80}	0.35	-0.09	-0.15	-0.20	0.30	0.35	-0.49	0.59	-0.03
P_{90}	0.35	0.01	-0.26	-0.14	-0.19	0.39	0.42	-0.15	-0.63
P_{95}	0.35	0.05	-0.31	-0.12	-0.21	0.01	-0.51	-0.63	0.23
P_{99}	0.35	0.07	-0.30	-0.09	-0.47	-0.57	0.18	0.42	0.15
P_{max}	0.35	0.15	-0.14	0.88	0.25	-0.02	0.03	0.00	-0.01

Note. Q_4, P_{mean}, P_{max} respectively refer to at-site 4th streamflow, at-site mean, and maximum daily precipitation. The rest of the variables represent percentiles of at-site daily precipitation.

Table S23. Mean of loadings and variance (in parenthesis) of all climatological variables on each principal component (PC) across all study sites.

Variables	PC1 (87%)	PC2 (10%)	PC3 (2%)	PC4 (1%)	PC5 (0%)	PC6 (0%)	PC7 (0%)	PC8 (0%)	PC9 (0%)
Q_5	0.19	0.89	0.36	-0.18	0.06	0.02	0.00	0.01	0.00
P_{mean}	0.32	-0.29	0.70	0.19	-0.49	0.16	-0.13	0.05	0.00
P_{70}	0.34	-0.24	0.27	-0.21	0.48	-0.56	-0.03	-0.19	-0.35
P_{75}	0.35	-0.16	0.07	-0.19	0.27	0.22	0.52	-0.10	0.64
P_{80}	0.35	-0.09	-0.15	-0.20	0.30	0.35	-0.49	0.59	-0.03
P_{90}	0.35	0.01	-0.26	-0.14	-0.19	0.39	0.42	-0.14	-0.63
P_{95}	0.35	0.05	-0.31	-0.12	-0.21	0.01	-0.51	-0.64	0.23
P_{99}	0.35	0.07	-0.30	-0.09	-0.47	-0.57	0.18	0.42	0.15
P_{max}	0.35	0.15	-0.14	0.88	0.25	-0.02	0.03	0.00	-0.01

Note. Q_5, P_{mean}, P_{max} respectively refer to at-site 5th streamflow, at-site mean, and maximum daily precipitation. The rest of the variables represent percentiles of at-site daily precipitation.

Table S24. Mean of loadings and variance (in parenthesis) of all climatological variables on each principal component (PC) across all study sites.

Variables	PC1 (87%)	PC2 (10%)	PC3 (2%)	PC4 (1%)	PC5 (0%)	PC6 (0%)	PC7 (0%)	PC8 (0%)	PC9 (0%)
Q_{10}	0.19	0.88	0.38	-0.19	0.05	0.02	0.00	0.01	0.00
P_{mean}	0.32	-0.30	0.69	0.18	-0.49	0.17	-0.13	0.05	0.00
P_{70}	0.34	-0.25	0.27	-0.21	0.47	-0.57	-0.03	-0.18	-0.36
P_{75}	0.35	-0.16	0.06	-0.19	0.27	0.22	0.52	-0.12	0.64
P_{80}	0.35	-0.09	-0.16	-0.21	0.30	0.35	-0.49	0.59	-0.01
P_{90}	0.35	0.01	-0.26	-0.14	-0.19	0.40	0.41	-0.14	-0.63
P_{95}	0.35	0.05	-0.31	-0.11	-0.21	-0.01	-0.52	-0.64	0.21
P_{99}	0.35	0.07	-0.30	-0.08	-0.47	-0.56	0.19	0.42	0.15
P_{max}	0.35	0.15	-0.13	0.88	0.26	-0.02	0.04	0.00	-0.01

Note. Q_{10}, P_{mean}, P_{max} respectively refer to at-site 10th streamflow, at-site mean, and maximum daily precipitation. The rest of the variables represent percentiles of at-site daily precipitation.

Table S25. Mean of loadings and variance (in parenthesis) of all climatological variables on each principal component (PC) across all study sites.

Variables	PC1 (88%)	PC2 (9%)	PC3 (2%)	PC4 (1%)	PC5 (0%)	PC6 (0%)	PC7 (0%)	PC8 (0%)	PC9 (0%)
Q_{20}	0.20	0.88	0.39	-0.19	0.04	0.02	0.00	0.01	0.00
P_{mean}	0.32	-0.32	0.69	0.20	-0.48	0.18	-0.14	0.05	0.00
P_{70}	0.34	-0.25	0.26	-0.22	0.45	-0.58	-0.02	-0.16	-0.36
P_{75}	0.35	-0.17	0.06	-0.20	0.27	0.21	0.52	-0.13	0.63
P_{80}	0.35	-0.08	-0.16	-0.22	0.31	0.36	-0.49	0.57	-0.01
P_{90}	0.35	0.01	-0.26	-0.14	-0.19	0.40	0.41	-0.15	-0.63
P_{95}	0.35	0.05	-0.31	-0.10	-0.22	-0.03	-0.52	-0.64	0.21
P_{99}	0.35	0.08	-0.31	-0.06	-0.48	-0.54	0.19	0.45	0.15
P_{max}	0.35	0.15	-0.13	0.87	0.28	-0.02	0.04	0.00	-0.01

Note. Q_{20}, P_{mean}, P_{max} respectively refer to at-site 20th streamflow, at-site mean, and maximum daily precipitation. The rest of the variables represent percentiles of at-site daily precipitation.

Table S26. Mean of loadings and variance (in parenthesis) of all climatological variables on each principal component (PC) across all study sites.

Variables	PC1 (88%)	PC2 (9%)	PC3 (2%)	PC4 (1%)	PC5 (0%)	PC6 (0%)	PC7 (0%)	PC8 (0%)	PC9 (0%)
Q_{30}	0.20	0.87	0.41	-0.18	0.04	0.02	0.00	0.00	0.00
P_{mean}	0.32	-0.33	0.68	0.20	-0.48	0.19	-0.14	0.05	0.00
P_{70}	0.34	-0.26	0.26	-0.22	0.44	-0.59	-0.02	-0.14	-0.36
P_{75}	0.35	-0.17	0.06	-0.20	0.27	0.21	0.52	-0.13	0.63
P_{80}	0.35	-0.08	-0.16	-0.23	0.31	0.37	-0.49	0.56	0.00
P_{90}	0.35	0.02	-0.26	-0.14	-0.18	0.40	0.40	-0.16	-0.63
P_{95}	0.35	0.05	-0.31	-0.10	-0.22	-0.05	-0.52	-0.63	0.22
P_{99}	0.35	0.08	-0.31	-0.05	-0.48	-0.52	0.20	0.46	0.15
P_{max}	0.35	0.15	-0.13	0.87	0.30	-0.02	0.04	0.00	-0.01

Note. Q_{30}, P_{mean}, P_{max} respectively refer to at-site 30th streamflow, at-site mean, and maximum daily precipitation. The rest of the variables represent percentiles of at-site daily precipitation.

Table S27. Mean of loadings and variance (in parenthesis) of all climatological variables on each principal component (PC) across all study sites.

Variables	PC1 (88%)	PC2 (9%)	PC3 (2%)	PC4 (1%)	PC5 (0%)	PC6 (0%)	PC7 (0%)	PC8 (0%)	PC9 (0%)
Q_{40}	0.20	0.87	0.41	-0.18	0.03	0.01	0.00	0.00	0.00
P_{mean}	0.32	-0.34	0.67	0.21	-0.48	0.19	-0.14	0.05	0.00
P_{70}	0.34	-0.26	0.26	-0.23	0.43	-0.60	-0.03	-0.14	-0.36
P_{75}	0.35	-0.17	0.06	-0.21	0.27	0.20	0.52	-0.13	0.63
P_{80}	0.35	-0.08	-0.16	-0.23	0.31	0.38	-0.49	0.56	-0.01
P_{90}	0.35	0.02	-0.26	-0.13	-0.18	0.40	0.41	-0.17	-0.63
P_{95}	0.35	0.06	-0.32	-0.09	-0.22	-0.05	-0.51	-0.63	0.22
P_{99}	0.35	0.08	-0.31	-0.04	-0.49	-0.51	0.19	0.47	0.15
P_{max}	0.34	0.15	-0.12	0.86	0.31	-0.02	0.04	0.00	-0.01

Note. Q_{40}, P_{mean}, P_{max} respectively refer to at-site 40th streamflow, at-site mean, and maximum daily precipitation. The rest of the variables represent percentiles of at-site daily precipitation.

Table S28. Mean of loadings and variance (in parenthesis) of all climatological variables on each principal component (PC) across all study sites.

Variables	PC1 (88%)	PC2 (9%)	PC3 (2%)	PC4 (1%)	PC5 (0%)	PC6 (0%)	PC7 (0%)	PC8 (0%)	PC9 (0%)
Q_{60}	0.21	0.86	0.43	-0.18	0.03	0.01	0.00	0.00	0.00
P_{mean}	0.32	-0.36	0.67	0.20	-0.47	0.19	-0.14	0.05	0.00
P_{70}	0.34	-0.26	0.25	-0.23	0.43	-0.60	-0.02	-0.14	-0.36
P_{75}	0.35	-0.17	0.05	-0.21	0.27	0.20	0.52	-0.13	0.63
P_{80}	0.35	-0.07	-0.17	-0.23	0.31	0.37	-0.49	0.56	0.00
P_{90}	0.35	0.02	-0.26	-0.13	-0.18	0.40	0.40	-0.17	-0.63
P_{95}	0.35	0.06	-0.32	-0.08	-0.22	-0.06	-0.51	-0.63	0.22
P_{99}	0.35	0.08	-0.31	-0.03	-0.49	-0.50	0.20	0.47	0.15
P_{max}	0.34	0.15	-0.11	0.86	0.31	-0.02	0.04	0.00	-0.01

Note. Q_{60}, P_{mean}, P_{max} respectively refer to at-site 60th streamflow, at-site mean, and maximum daily precipitation. The rest of the variables represent percentiles of at-site daily precipitation.

Table S29. Mean of loadings and variance (in parenthesis) of all climatological variables on each principal component (PC) across all study sites.

Variables	PC1 (88%)	PC2 (9%)	PC3 (2%)	PC4 (1%)	PC5 (0%)	PC6 (0%)	PC7 (0%)	PC8 (0%)	PC9 (0%)
Q_{70}	0.22	0.84	0.45	-0.20	0.04	0.01	0.00	0.00	0.00
P_{mean}	0.32	-0.38	0.66	0.17	-0.47	0.19	-0.14	0.05	0.01
P_{70}	0.34	-0.28	0.23	-0.24	0.44	-0.60	-0.02	-0.14	-0.36
P_{75}	0.35	-0.18	0.03	-0.21	0.27	0.20	0.53	-0.13	0.63
P_{80}	0.35	-0.07	-0.18	-0.22	0.31	0.38	-0.49	0.55	0.00
P_{90}	0.35	0.03	-0.27	-0.12	-0.18	0.40	0.40	-0.17	-0.64
P_{95}	0.35	0.07	-0.32	-0.07	-0.23	-0.06	-0.51	-0.63	0.23
P_{99}	0.35	0.09	-0.30	-0.03	-0.49	-0.50	0.20	0.47	0.14
P_{max}	0.34	0.15	-0.08	0.87	0.31	-0.02	0.04	0.00	-0.01

Note. Q_{70}, P_{mean}, P_{max} respectively refer to at-site 70th streamflow, at-site mean, and maximum daily precipitation. The rest of the variables represent percentiles of at-site daily precipitation.

Table S30. Mean of loadings and variance (in parenthesis) of all climatological variables on each principal component (PC) across all study sites.

Variables	PC1 (88%)	PC2 (9%)	PC3 (2%)	PC4 (1%)	PC5 (0%)	PC6 (0%)	PC7 (0%)	PC8 (0%)	PC9 (0%)
Q_{80}	0.22	0.84	0.45	-0.20	0.04	0.01	0.00	0.00	0.00
P_{mean}	0.32	-0.38	0.67	0.18	-0.47	0.19	-0.14	0.05	0.01
P_{70}	0.34	-0.28	0.23	-0.24	0.44	-0.60	-0.02	-0.14	-0.36
P_{75}	0.35	-0.17	0.03	-0.21	0.27	0.20	0.53	-0.13	0.63
P_{80}	0.35	-0.07	-0.18	-0.22	0.31	0.37	-0.50	0.55	0.00
P_{90}	0.35	0.03	-0.27	-0.12	-0.18	0.41	0.40	-0.17	-0.64
P_{95}	0.35	0.07	-0.32	-0.08	-0.23	-0.06	-0.51	-0.63	0.23
P_{99}	0.35	0.09	-0.30	-0.03	-0.49	-0.50	0.20	0.47	0.14
P_{max}	0.34	0.15	-0.08	0.87	0.31	-0.02	0.04	0.00	-0.01

Note. Q_{80}, P_{mean}, P_{max} respectively refer to at-site 80th streamflow, at-site mean, and maximum daily precipitation. The rest of the variables represent percentiles of at-site daily precipitation.

Table S31. Mean of loadings and variance (in parenthesis) of all climatological variables on each principal component (PC) across all study sites.

Variables	PC1 (88%)	PC2 (9%)	PC3 (2%)	PC4 (1%)	PC5 (0%)	PC6 (0%)	PC7 (0%)	PC8 (0%)	PC9 (0%)
Q_{90}	0.22	0.84	0.45	-0.20	0.04	0.01	0.00	0.00	0.00
P_{mean}	0.32	-0.38	0.66	0.18	-0.47	0.19	-0.14	0.05	0.01
P_{70}	0.34	-0.28	0.23	-0.24	0.44	-0.60	-0.03	-0.14	-0.36
P_{75}	0.35	-0.17	0.04	-0.21	0.27	0.20	0.53	-0.13	0.63
P_{80}	0.35	-0.07	-0.18	-0.23	0.31	0.38	-0.49	0.55	0.00
P_{90}	0.35	0.03	-0.27	-0.12	-0.18	0.40	0.40	-0.17	-0.64
P_{95}	0.35	0.07	-0.32	-0.08	-0.23	-0.06	-0.51	-0.63	0.23
P_{99}	0.35	0.09	-0.30	-0.03	-0.49	-0.50	0.20	0.47	0.14
P_{max}	0.34	0.15	-0.09	0.87	0.31	-0.02	0.04	0.00	-0.01

Note. Q_{90}, P_{mean}, P_{max} respectively refer to at-site 90th streamflow, at-site mean, and maximum daily precipitation. The rest of the variables represent percentiles of at-site daily precipitation.

Table S32. Mean of loadings and variance (in parenthesis) of all climatological variables on each principal component (PC) across all study sites.

Variables	PC1 (88%)	PC2 (9%)	PC3 (2%)	PC4 (1%)	PC5 (0%)	PC6 (0%)	PC7 (0%)	PC8 (0%)	PC9 (0%)
Q_{95}	0.22	0.84	0.45	-0.19	0.03	0.01	0.00	0.00	0.00
P_{mean}	0.32	-0.38	0.66	0.19	-0.47	0.19	-0.14	0.05	0.01
P_{70}	0.34	-0.28	0.23	-0.24	0.43	-0.60	-0.03	-0.13	-0.36
P_{75}	0.35	-0.17	0.04	-0.21	0.27	0.20	0.53	-0.13	0.63
P_{80}	0.35	-0.07	-0.18	-0.23	0.31	0.38	-0.49	0.55	0.00
P_{90}	0.35	0.03	-0.27	-0.13	-0.18	0.40	0.40	-0.17	-0.63
P_{95}	0.35	0.07	-0.32	-0.08	-0.23	-0.06	-0.51	-0.63	0.22
P_{99}	0.35	0.09	-0.30	-0.03	-0.49	-0.51	0.19	0.47	0.14
P_{max}	0.34	0.15	-0.09	0.87	0.31	-0.02	0.04	0.00	-0.01

Note. Q_{95}, P_{mean}, P_{max} respectively refer to at-site 95th streamflow, at-site mean, and maximum daily precipitation. The rest of the variables represent percentiles of at-site daily precipitation.

Table S33. Mean of loadings and variance (in parenthesis) of all climatological variables on each principal component (PC) across all study sites.

Variables	PC1 (89%)	PC2 (9%)	PC3 (2%)	PC4 (1%)	PC5 (0%)	PC6 (0%)	PC7 (0%)	PC8 (0%)	PC9 (0%)
Q_{96}	0.22	0.84	0.45	-0.19	0.03	0.01	0.00	0.00	0.00
P_{mean}	0.32	-0.38	0.66	0.19	-0.47	0.19	-0.14	0.05	0.01
P_{70}	0.34	-0.28	0.24	-0.24	0.43	-0.60	-0.03	-0.13	-0.36
P_{75}	0.35	-0.17	0.04	-0.21	0.27	0.19	0.53	-0.13	0.63
P_{80}	0.35	-0.07	-0.17	-0.23	0.31	0.38	-0.49	0.55	0.00
P_{90}	0.35	0.03	-0.27	-0.13	-0.18	0.40	0.40	-0.17	-0.63
P_{95}	0.35	0.07	-0.32	-0.08	-0.23	-0.06	-0.51	-0.64	0.22
P_{99}	0.35	0.09	-0.30	-0.03	-0.49	-0.50	0.19	0.47	0.15
P_{max}	0.34	0.15	-0.10	0.87	0.31	-0.02	0.04	0.00	-0.01

Note. Q_{96}, P_{mean}, P_{max} respectively refer to at-site 96^h streamflow, at-site mean, and maximum daily precipitation. The rest of the variables represent percentiles of at-site daily precipitation.

Table S34. Mean of loadings and variance (in parenthesis) of all climatological variables on each principal component (PC) across all study sites.

Variables	PC1 (89%)	PC2 (9%)	PC3 (2%)	PC4 (1%)	PC5 (0%)	PC6 (0%)	PC7 (0%)	PC8 (0%)	PC9 (0%)
Q_{97}	0.22	0.84	0.45	-0.19	0.03	0.01	0.00	0.00	0.00
P_{mean}	0.32	-0.38	0.66	0.20	-0.47	0.19	-0.14	0.04	0.01
P_{70}	0.34	-0.28	0.24	-0.24	0.43	-0.60	-0.03	-0.13	-0.36
P_{75}	0.35	-0.17	0.04	-0.21	0.27	0.19	0.53	-0.13	0.63
P_{80}	0.35	-0.07	-0.17	-0.23	0.31	0.38	-0.49	0.55	0.00
P_{90}	0.35	0.03	-0.27	-0.13	-0.18	0.40	0.41	-0.17	-0.63
P_{95}	0.35	0.07	-0.32	-0.08	-0.23	-0.05	-0.51	-0.64	0.22
P_{99}	0.35	0.09	-0.31	-0.03	-0.49	-0.51	0.19	0.48	0.14
P_{max}	0.34	0.15	-0.10	0.87	0.32	-0.02	0.04	0.00	-0.01

Note. Q_{97}, P_{mean}, P_{max} respectively refer to at-site 97^h streamflow, at-site mean, and maximum daily precipitation. The rest of the variables represent percentiles of at-site daily precipitation.

Table S35. Mean of loadings and variance (in parenthesis) of all climatological variables on each principal component (PC) across all study sites.

Variables	PC1 (89%)	PC2 (9%)	PC3 (2%)	PC4 (1%)	PC5 (0%)	PC6 (0%)	PC7 (0%)	PC8 (0%)	PC9 (0%)
Q_{98}	0.22	0.84	0.45	-0.19	0.03	0.01	0.00	0.00	0.00
P_{mean}	0.32	-0.38	0.66	0.20	-0.47	0.20	-0.14	0.04	0.01
P_{70}	0.34	-0.27	0.24	-0.24	0.43	-0.60	-0.03	-0.13	-0.36
P_{75}	0.35	-0.17	0.04	-0.21	0.27	0.19	0.53	-0.13	0.63
P_{80}	0.35	-0.07	-0.17	-0.23	0.31	0.39	-0.49	0.55	0.00
P_{90}	0.35	0.03	-0.27	-0.13	-0.18	0.40	0.41	-0.17	-0.63
P_{95}	0.35	0.07	-0.32	-0.08	-0.23	-0.05	-0.51	-0.64	0.22
P_{99}	0.35	0.09	-0.31	-0.03	-0.49	-0.50	0.18	0.48	0.14
P_{max}	0.34	0.14	-0.10	0.86	0.32	-0.02	0.04	0.00	-0.01

Note. Q_{98}, P_{mean}, P_{max} respectively refer to at-site 98^h streamflow, at-site mean, and maximum daily precipitation. The rest of the variables represent percentiles of at-site daily precipitation.

Table S36. Mean of loadings and variance (in parenthesis) of all climatological variables on each principal component (PC) across all study sites.

Variables	PC1 (89%)	PC2 (9%)	PC3 (2%)	PC4 (1%)	PC5 (0%)	PC6 (0%)	PC7 (0%)	PC8 (0%)	PC9 (0%)
Q_{99}	0.22	0.84	0.45	-0.18	0.03	0.01	0.00	0.00	0.00
P_{mean}	0.32	-0.38	0.65	0.21	-0.47	0.20	-0.13	0.04	0.01
P_{70}	0.34	-0.27	0.24	-0.24	0.43	-0.60	-0.04	-0.13	-0.36
P_{75}	0.35	-0.17	0.04	-0.21	0.27	0.18	0.53	-0.13	0.63
P_{80}	0.35	-0.07	-0.17	-0.24	0.31	0.39	-0.49	0.55	-0.01
P_{90}	0.35	0.03	-0.26	-0.13	-0.18	0.39	0.42	-0.17	-0.63
P_{95}	0.35	0.07	-0.32	-0.08	-0.22	-0.05	-0.51	-0.64	0.22
P_{99}	0.35	0.09	-0.31	-0.03	-0.49	-0.51	0.18	0.48	0.14
P_{max}	0.34	0.14	-0.11	0.86	0.32	-0.02	0.04	0.00	-0.01

Note. Q_{99}, P_{mean}, P_{max} respectively refer to at-site 99^h streamflow, at-site mean, and maximum daily precipitation. The rest of the variables represent percentiles of at-site daily precipitation.

Table S37. Mean of loadings and variance (in parenthesis) of all climatological, hydrological, and land cover variables on each principal component (PC) across all study sites.

Variables	PC1 (51%)	PC2 (14%)	PC3 (11%)	PC4 (6%)	PC5 (5%)	PC6 (5%)	PC7 (3%)	PC8 (2%)	PC9 (2%)	PC10 (1%)	PC11 (0%)	PC12 (0%)	PC13 (0%)	PC14 (0%)	PC15 (0%)	PC16 (0%)	PC17 (0%)	PC18 (0%)	PC19 (0%)
Q_{mean}	0.22	-0.09	0.33	0.41	0.20	0.05	0.21	-0.09	0.04	-0.06	0.17	0.18	0.01	-0.01	0.60	0.21	0.28	0.14	0.02
P_{mean}	0.28	0.18	-0.22	-0.11	-0.03	-0.03	-0.09	-0.08	0.14	-0.10	0.61	-0.18	0.27	-0.49	0.13	-0.19	-0.12	-0.01	-0.02
P_{70}	0.30	0.10	-0.16	-0.18	0.00	-0.04	-0.01	-0.03	0.11	0.03	0.33	-0.04	-0.09	0.44	-0.28	0.19	0.51	0.05	0.37
P_{75}	0.31	0.08	-0.10	-0.17	0.02	-0.05	0.03	-0.02	0.04	0.02	0.13	0.09	-0.13	0.29	0.07	0.42	-0.29	-0.25	-0.63
P_{80}	0.31	0.04	-0.05	-0.18	0.04	-0.05	0.07	0.00	-0.02	0.01	-0.06	0.10	-0.18	0.35	0.21	-0.57	-0.26	0.50	-0.01
P_{90}	0.31	0.04	0.01	-0.15	0.06	-0.06	0.09	0.00	-0.10	-0.01	-0.17	0.22	-0.11	-0.14	0.10	0.22	-0.45	-0.29	0.63
P_{95}	0.31	0.03	0.03	-0.13	0.06	-0.06	0.08	0.00	-0.12	-0.01	-0.23	0.16	-0.10	-0.16	0.02	-0.45	0.45	-0.55	-0.18
P_{99}	0.31	0.03	0.04	-0.12	0.06	-0.07	0.08	-0.01	-0.13	-0.03	-0.24	0.17	-0.09	-0.49	-0.36	0.27	0.15	0.51	-0.17
P_{max}	0.31	0.10	0.06	0.00	0.03	-0.05	0.11	0.05	-0.12	-0.09	-0.39	-0.65	0.48	0.15	0.09	0.09	0.01	0.02	0.01
WAT	-0.04	0.43	-0.23	0.21	-0.33	0.17	0.26	-0.68	0.05	0.14	-0.17	0.04	-0.04	0.00	0.01	0.00	0.01	0.00	0.00
BUL	-0.20	0.25	0.33	-0.33	-0.06	-0.26	0.24	-0.01	0.13	-0.04	0.03	0.43	0.58	0.12	-0.05	-0.01	0.01	0.02	-0.01
OPN	-0.13	0.12	-0.09	-0.15	0.82	0.11	-0.18	-0.30	0.32	0.07	-0.15	-0.02	0.03	-0.01	-0.01	0.00	0.00	0.00	0.00
VEG	-0.25	-0.16	-0.15	-0.22	0.10	-0.12	0.50	-0.14	-0.06	-0.68	0.07	-0.16	-0.22	-0.01	0.03	0.02	0.02	0.00	0.01
AGR	-0.20	0.34	0.13	-0.13	0.26	0.03	0.21	0.06	-0.68	0.35	0.27	-0.15	-0.12	-0.02	0.03	0.01	0.01	0.00	0.00
WTL	-0.04	0.44	-0.24	0.22	0.09	0.19	0.37	0.63	0.32	-0.01	-0.08	0.04	-0.08	-0.05	-0.02	-0.01	0.00	-0.01	0.00
Imp	-0.04	0.26	0.55	-0.15	-0.14	-0.30	-0.04	-0.03	0.39	0.08	-0.01	-0.35	-0.44	-0.12	0.05	0.00	0.00	-0.02	0.01
A	0.21	-0.10	0.35	0.44	0.19	0.05	0.24	-0.13	0.02	-0.08	0.21	-0.03	0.04	0.10	-0.58	-0.20	-0.26	-0.13	-0.03
S	0.02	-0.49	0.00	-0.31	-0.06	0.28	0.48	-0.02	0.25	0.51	0.04	-0.13	0.06	-0.11	0.02	0.00	0.00	0.00	0.00
DD	0.05	0.14	0.34	-0.28	-0.12	0.80	-0.16	0.02	-0.04	-0.31	0.01	0.02	-0.01	0.02	0.00	0.01	0.01	0.00	-0.01

Table S38. Mean of loadings and variance (in parenthesis) of all climatological, hydrological, and land cover variables on each principal component (PC) across all study sites.

Variables	PC1 (51%)	PC2 (14%)	PC3 (11%)	PC4 (6%)	PC5 (5%)	PC6 (5%)	PC7 (3%)	PC8 (2%)	PC9 (2%)	PC10 (1%)	PC11 (0%)	PC12 (0%)	PC13 (0%)	PC14 (0%)	PC15 (0%)	PC16 (0%)	PC17 (0%)	PC18 (0%)	PC19 (0%)
Q_{min}	0.14	-0.08	0.41	0.49	0.15	-0.05	0.12	-0.11	0.07	-0.54	0.23	-0.40	0.01	-0.01	-0.01	0.00	0.01	0.01	0.00
P_{mean}	0.28	0.18	-0.21	-0.04	-0.01	0.02	-0.13	-0.09	0.15	-0.37	0.23	0.44	-0.17	0.28	-0.49	-0.25	-0.04	0.04	-0.01
P_{70}	0.31	0.09	-0.16	-0.12	0.00	0.03	-0.04	-0.04	0.12	-0.20	0.21	0.17	-0.04	-0.09	0.44	0.56	-0.27	-0.15	0.34
P_{75}	0.31	0.07	-0.11	-0.13	0.03	0.04	0.02	-0.03	0.05	-0.11	0.10	0.04	0.09	-0.13	0.29	0.02	0.57	-0.13	-0.62
P_{80}	0.32	0.04	-0.06	-0.15	0.05	0.05	0.07	-0.01	-0.02	-0.02	0.01	-0.08	0.10	-0.18	0.35	-0.51	-0.29	0.58	0.02
P_{90}	0.32	0.03	0.00	-0.14	0.06	0.05	0.09	-0.01	-0.09	0.05	-0.08	-0.13	0.22	-0.11	-0.14	-0.18	0.51	-0.16	0.65
P_{95}	0.32	0.03	0.02	-0.13	0.07	0.06	0.10	0.00	-0.12	0.07	-0.11	-0.18	0.17	-0.10	-0.16	-0.20	-0.49	-0.63	-0.21
P_{99}	0.32	0.03	0.03	-0.12	0.07	0.07	0.09	-0.01	-0.13	0.08	-0.14	-0.16	0.18	-0.09	-0.49	0.54	-0.05	0.44	-0.17
P_{max}	0.31	0.10	0.06	-0.02	0.04	0.05	0.14	0.04	-0.13	0.14	-0.27	-0.24	-0.67	0.48	0.15	0.03	0.06	0.00	0.01
WAT	-0.04	0.42	-0.21	0.23	-0.37	-0.15	0.20	-0.68	0.03	0.18	0.02	-0.14	0.04	-0.04	0.00	0.00	-0.01	0.00	0.00
BUL	-0.20	0.26	0.27	-0.37	-0.06	0.26	0.20	-0.03	0.14	-0.13	0.02	-0.06	0.43	0.58	0.12	0.03	-0.02	0.02	-0.01
OPN	-0.13	0.13	-0.12	-0.10	0.83	-0.14	-0.15	-0.30	0.30	0.14	-0.03	-0.11	-0.02	0.03	-0.01	0.00	0.00	0.00	0.00
VEG	-0.25	-0.15	-0.19	-0.19	0.07	0.12	0.46	-0.19	-0.04	-0.50	-0.48	0.11	-0.16	-0.22	-0.01	0.01	0.00	-0.02	0.00
AGR	-0.20	0.34	0.09	-0.14	0.24	-0.04	0.23	0.06	-0.67	-0.01	0.45	0.09	-0.15	-0.12	-0.02	-0.01	0.00	0.00	0.00
WTL	-0.03	0.44	-0.23	0.24	0.04	-0.18	0.41	0.60	0.34	0.03	-0.05	-0.05	0.05	-0.08	-0.05	0.00	-0.01	0.00	0.00
Imp	-0.04	0.27	0.51	-0.24	-0.12	0.30	-0.05	-0.04	0.39	0.10	0.04	0.04	-0.35	-0.44	-0.12	-0.03	0.01	-0.02	0.01
A	0.20	-0.09	0.38	0.31	0.18	-0.05	0.36	-0.13	0.00	0.27	-0.15	0.63	0.13	0.06	0.09	0.01	-0.01	0.00	-0.01
S	0.02	-0.48	-0.04	-0.31	-0.08	-0.27	0.47	-0.06	0.26	0.19	0.46	-0.11	-0.13	0.05	-0.11	-0.01	0.00	0.00	0.00
DD	0.05	0.14	0.31	-0.30	-0.12	-0.81	-0.15	0.02	-0.04	-0.19	-0.24	0.05	0.02	-0.01	0.02	0.01	0.00	0.00	-0.01

Table S39. Mean of loadings and variance (in parenthesis) of all climatological, hydrological, and land cover variables on each principal component (PC) across all study sites.

Variables	PC1 (51%)	PC2 (14%)	PC3 (11%)	PC4 (6%)	PC5 (5%)	PC6 (5%)	PC7 (3%)	PC8 (2%)	PC9 (2%)	PC10 (1%)	PC11 (0%)	PC12 (0%)	PC13 (0%)	PC14 (0%)	PC15 (0%)	PC16 (0%)	PC17 (0%)	PC18 (0%)	PC19 (0%)
Q_{50}	0.21	-0.10	0.34	0.43	0.19	0.05	0.20	-0.10	0.05	-0.10	0.22	0.14	0.03	0.32	-0.60	0.06	0.01	0.01	0.00
P_{mean}	0.28	0.18	-0.22	-0.10	-0.02	-0.03	-0.09	-0.08	0.14	-0.14	0.59	-0.20	0.27	-0.46	-0.19	-0.25	-0.04	0.04	-0.01
P_{70}	0.30	0.10	-0.16	-0.17	0.00	-0.04	-0.01	-0.03	0.11	0.01	0.33	-0.04	-0.08	0.37	0.28	0.52	-0.27	-0.15	0.34
P_{75}	0.31	0.08	-0.10	-0.17	0.02	-0.05	0.04	-0.02	0.04	0.01	0.14	0.09	-0.13	0.27	0.07	0.09	0.57	-0.13	-0.62
P_{80}	0.31	0.04	-0.05	-0.17	0.05	-0.05	0.07	0.00	-0.02	0.01	-0.05	0.11	-0.18	0.32	0.15	-0.57	-0.29	0.58	0.02
P_{90}	0.31	0.04	0.01	-0.14	0.06	-0.06	0.09	0.00	-0.10	0.00	-0.17	0.22	-0.12	-0.13	-0.10	-0.14	0.51	-0.16	0.65
P_{95}	0.31	0.03	0.03	-0.13	0.07	-0.06	0.08	0.00	-0.12	0.00	-0.23	0.17	-0.10	-0.14	-0.08	-0.19	-0.49	-0.63	-0.21
P_{99}	0.31	0.03	0.04	-0.12	0.07	-0.07	0.08	0.00	-0.14	-0.02	-0.24	0.17	-0.10	-0.46	-0.13	0.52	-0.05	0.44	-0.17
P_{max}	0.31	0.10	0.06	0.00	0.04	-0.05	0.11	0.05	-0.13	-0.06	-0.40	-0.64	0.49	0.19	-0.01	0.04	0.06	0.00	0.01
WAT	-0.04	0.43	-0.23	0.21	-0.33	0.17	0.26	-0.68	0.03	0.15	-0.16	0.04	-0.04	0.01	-0.02	0.00	-0.01	0.00	0.00
BUL	-0.20	0.25	0.32	-0.33	-0.05	-0.26	0.25	-0.01	0.13	-0.05	0.04	0.44	0.57	0.07	0.09	0.02	-0.02	0.02	-0.01
OPN	-0.13	0.12	-0.09	-0.13	0.83	0.12	-0.18	-0.30	0.31	0.08	-0.15	-0.02	0.04	0.00	0.00	0.00	0.00	0.00	0.00
VEG	-0.25	-0.16	-0.16	-0.21	0.11	-0.12	0.50	-0.14	-0.06	-0.68	0.02	-0.17	-0.22	0.00	-0.02	0.01	0.00	-0.02	0.00
AGR	-0.20	0.34	0.13	-0.12	0.26	0.04	0.21	0.07	-0.68	0.33	0.30	-0.15	-0.12	0.00	-0.04	0.00	0.00	0.00	0.00
WTL	-0.04	0.44	-0.23	0.23	0.08	0.19	0.37	0.62	0.33	0.00	-0.09	0.04	-0.08	-0.05	-0.01	0.00	-0.01	0.00	0.00
Imp	-0.04	0.26	0.54	-0.17	-0.14	-0.30	-0.04	-0.04	0.39	0.09	-0.01	-0.36	-0.44	-0.08	-0.08	-0.02	0.01	-0.02	0.01
A	0.21	-0.10	0.35	0.42	0.18	0.06	0.23	-0.12	0.01	-0.06	0.13	0.00	0.01	-0.26	0.66	-0.06	-0.01	0.00	-0.01
S	0.02	-0.49	-0.01	-0.30	-0.05	0.28	0.49	-0.02	0.25	0.50	0.07	-0.13	0.06	-0.08	-0.07	-0.01	0.00	0.00	0.00
DD	0.05	0.14	0.33	-0.29	-0.12	0.80	-0.15	0.02	-0.04	-0.31	-0.01	0.02	-0.01	0.02	0.01	0.01	0.00	0.00	-0.01

Table S40. Mean of loadings and variance (in parenthesis) of all climatological, hydrological, and land cover variables on each principal component (PC) across all study sites.

Variables	PC1 (51%)	PC2 (14%)	PC3 (11%)	PC4 (6%)	PC5 (5%)	PC6 (5%)	PC7 (3%)	PC8 (2%)	PC9 (2%)	PC10 (1%)	PC11 (0%)	PC12 (0%)	PC13 (0%)	PC14 (0%)	PC15 (0%)	PC16 (0%)	PC17 (0%)	PC18 (0%)	PC19 (0%)
Q_{max}	0.20	-0.10	0.33	0.47	0.18	0.05	0.21	-0.11	0.07	-0.05	0.25	0.57	0.29	0.20	-0.01	-0.03	-0.01	-0.02	0.00
P_{mean}	0.28	0.18	-0.22	-0.10	0.00	-0.02	-0.08	-0.10	0.15	-0.10	0.60	0.03	-0.11	-0.31	-0.49	0.25	0.07	0.06	-0.01
P_{70}	0.30	0.10	-0.16	-0.17	0.03	-0.03	0.00	-0.04	0.12	0.03	0.33	-0.03	-0.07	0.07	0.44	-0.58	0.20	-0.12	0.35
P_{75}	0.31	0.08	-0.10	-0.16	0.05	-0.04	0.04	-0.03	0.04	0.02	0.13	-0.07	0.04	0.14	0.29	0.06	-0.52	-0.24	-0.63
P_{80}	0.31	0.05	-0.04	-0.17	0.07	-0.05	0.07	0.00	-0.02	0.01	-0.05	-0.02	0.07	0.22	0.35	0.43	0.27	0.65	0.00
P_{90}	0.31	0.04	0.01	-0.14	0.07	-0.06	0.08	0.01	-0.10	-0.01	-0.17	-0.10	0.17	0.15	-0.15	0.26	-0.47	-0.20	0.64
P_{95}	0.31	0.03	0.03	-0.12	0.08	-0.07	0.08	0.02	-0.13	-0.01	-0.23	-0.06	0.13	0.13	-0.16	0.14	0.62	-0.55	-0.17
P_{99}	0.31	0.03	0.04	-0.11	0.07	-0.07	0.07	0.01	-0.14	-0.03	-0.24	-0.09	0.13	0.11	-0.49	-0.56	-0.09	0.39	-0.19
P_{max}	0.31	0.10	0.06	0.01	0.03	-0.05	0.11	0.06	-0.13	-0.10	-0.36	0.43	-0.44	-0.56	0.16	-0.01	-0.06	0.00	0.01
WAT	-0.04	0.43	-0.23	0.18	-0.32	0.19	0.28	-0.68	-0.03	0.14	-0.17	0.01	0.03	0.05	0.00	-0.01	0.01	0.00	0.00
BUL	-0.20	0.25	0.33	-0.33	-0.01	-0.25	0.26	-0.02	0.13	-0.04	0.03	-0.04	0.56	-0.46	0.12	-0.03	0.01	0.02	-0.01
OPN	-0.13	0.12	-0.08	-0.02	0.84	0.07	-0.20	-0.32	0.27	0.07	-0.18	-0.03	-0.03	-0.05	0.00	0.00	0.00	0.00	0.00
VEG	-0.25	-0.16	-0.16	-0.20	0.15	-0.11	0.50	-0.13	-0.08	-0.68	0.07	0.02	-0.21	0.18	-0.01	-0.01	0.00	-0.02	0.01
AGR	-0.20	0.34	0.13	-0.09	0.28	0.02	0.19	0.13	-0.67	0.35	0.29	0.04	-0.17	0.09	-0.02	0.01	0.00	0.00	0.00
WTL	-0.04	0.44	-0.23	0.25	0.08	0.19	0.36	0.60	0.37	-0.01	-0.10	-0.06	0.00	0.07	-0.05	0.00	0.01	0.00	0.00
Imp	-0.04	0.26	0.54	-0.17	-0.14	-0.29	-0.02	-0.08	0.39	0.08	-0.02	0.03	-0.45	0.34	-0.12	0.03	-0.01	-0.02	0.01
A	0.21	-0.10	0.35	0.45	0.10	0.03	0.22	-0.10	-0.05	-0.09	0.10	-0.67	-0.16	-0.22	0.10	0.02	0.02	0.02	0.00
S	0.02	-0.49	0.00	-0.29	0.02	0.29	0.49	-0.04	0.23	0.51	0.03	0.03	-0.12	-0.09	-0.10	0.01	0.00	0.00	0.00
DD	0.05	0.14	0.34	-0.28	-0.05	0.81	-0.16	0.02	-0.03	-0.31	0.02	0.00	0.02	0.02	0.02	-0.01	0.00	0.00	-0.01

Table S41. Features of 36 flow stations across the Monongahela River Bain USA.

USGS ID	Name	State	LAT(°N)	LONG(°w)	Start Date	Finish Date
03083500	Youghiogheny River at Sutersville	PA	40.24	-79.81	01/01/1990	31/12/2020
03061000	West fork river at enterprise	WV	39.42	-80.28	01/01/1990	31/12/2020
03081500	Youghiogheny River at Ohiopyle	PA	39.87	-79.49	01/10/2002	31/12/2020
03067510	Shavers fork near cheat bridge	WV	38.62	-79.87	15/02/2000	31/12/2020
03085000	Monongahela River at Braddock	PA	40.39	-79.86	01/01/1990	26/02/2017
03075070	Monongahela River at Eliz-abeth	PA	40.26	-79.90	01/01/1990	31/12/2020
03082500	Youghiogheny River at Connellsville	PA	40.02	-79.59	01/01/1990	31/12/2020
03074500	Redstone Creek at Walters-burg	PA	39.98	-79.76	01/01/1990	31/12/2020
03073000	South Fork Tenmile Creek at Jefferson	PA	39.92	-80.07	01/01/1990	31/12/2020
03079000	Casselman River at Mar-kleton	PA	39.86	-79.23	01/01/1990	31/12/2020
03081000	Youghiogheny River below Confluence	PA	39.83	-79.37	01/01/1990	31/12/2020
03072655	Monongahela River near Masontown	PA	39.83	-79.92	01/01/1990	31/12/2020
03080000	Laurel Hill Creek at Ursina	PA	39.82	-79.32	01/01/1990	31/12/2020
03077500	Youghiogheny River at Youghiogheny River Dam	PA	39.81	-79.36	01/01/1990	31/12/2020
03072000	Dunkard Creek at Shan-nopin	PA	39.76	-79.97	01/01/1990	31/12/2020
03076600	Bear creek at Friendsville	MD	39.66	-79.39	01/01/1990	31/12/2020
03076500	Youghiogheny river at Friendsville	MD	39.65	-79.41	01/01/1990	31/12/2020
03070500	Big sandy creek at Rock-ville	WV	39.62	-79.70	01/01/1990	31/12/2020
03076100	Youghiogheny river at Hoyes	MD	39.53	-79.41	29/07/2011	31/12/2020
03070260	Cheat river at Albright	WV	39.49	-79.64	01/10/1996	31/12/2020
03061500	Buffalo creek at barrack Ville	WV	39.51	-80.17	01/01/1990	31/12/2020
03057000	Tygart valley river at Colfax	WV	39.44	-80.13	01/01/1990	31/12/2020
03075500	Youghiogheny river near Oakland	MD	39.42	-79.42	01/01/1990	31/12/2020
03056250	Three fork creek near Graf-ton	WV	39.34	-79.99	01/01/1990	31/12/2020
03056000	Tygart valley r at tygart dam nr grafton	WV	39.32	-80.03	01/01/1990	31/12/2020
03054500	Tygart valley river at Phi-lippi	WV	39.15	-80.04	01/01/1990	31/12/2020
03069500	Cheat river near parsons	WV	39.12	-79.68	01/01/1990	31/12/2020
03065000	Dry fork at Hendricks	WV	39.07	-79.62	01/01/1990	31/12/2020
03053500	Buckhannon river at hall	WV	39.05	-80.11	01/01/1990	31/12/2020
03052000	Middle fork river at Audra	WV	39.04	-80.07	01/01/1990	31/12/2020

Table S41. Continued

USGS ID	Name	State	LAT(°N)	LONG(°w)	Start Date	Finish Date
03051000	Tygart valley river at Belington	WV	39.03	-79.94	01/01/1990	31/12/2020
03052500	Sand run near Buckhannon	WV	38.96	-80.15	01/01/1990	31/12/2020
03068800	Shavers fork below Bowden	WV	38.91	-79.77	01/10/1997	31/12/2020
03052120	Buckhannon river at Alton	WV	38.82	-80.21	01/10/2011	31/12/2020
03050000	Tygart valley river near Dailey	WV	38.81	-79.88	01/01/1990	31/12/2020
03062500	Deckers creek at Morgantown	WV	39.63	-79.95	01/10/2002	31/12/2020

Table S42. Features of 47 climate stations across the Monongahela River Bain USA.

Station ID	Name	State	LAT(°N)	LONG(°w)	Start Date	Finish Date
USC00468308	SNOWSHOE	WV	38.40	-79.99	01/02/1975	31/12/2020
USC00467649	ROCK CAVE 2 NE	WV	38.86	-80.31	01/01/1953	31/12/2020
USC00468522	STONEWALL JACK-SON DAM	WV	39.00	-80.47	01/10/1999	31/12/2020
USC00464971	KUMBRABOW	WV	38.63	-80.09	01/11/1939	31/12/2020
USC00460633	BELINGTON	WV	39.02	-79.93	01/03/1938	14/03/2011
USC00469436	WESTON	WV	39.04	-80.47	01/01/1900	29/02/2012
USC00461220	BUCKHANNON	WV	38.98	-80.22	01/01/1900	31/12/2020
USC00466989	PICKENS 2 N	WV	38.68	-80.20	01/12/1994	01/10/2011
USC00461677	CLARKSBURG 1	WV	39.27	-80.35	17/11/1922	31/12/2020
USC00468986	TYGART DAM	WV	39.31	-80.03	01/08/1948	31/12/2020
USC00463464	GLADY 1 N	WV	38.81	-79.72	01/09/1973	31/08/2015
USC00466982	PHILIPPI	WV	39.15	-80.04	01/01/1900	31/12/2010
USC00469086	VALLEY HEAD	WV	38.55	-80.04	19/02/1938	20/04/2013
USW00013729	ELKINS RANDOLPH CO AIRPORT	WV	38.89	-79.86	01/01/1926	31/12/2020
USC00461397	CANAAN VALLEY NUMBER 2	WV	39.05	-79.44	01/01/1993	31/12/2020
USC00186620	OAKLAND 1 SE	MD	39.41	-79.40	01/01/1900	31/12/2020
USC00468777	TERRA ALTA NUMBER 1	WV	39.45	-79.55	01/08/1948	07/01/2020
USC00462211	DAVIS 3 SE	WV	39.10	-79.43	01/07/2002	31/12/2020
USC00462920	FAIRMONT	WV	39.46	-80.14	01/01/1900	31/12/2020
USW00003733	ELKINS 21 ENE	WV	39.01	-79.47	18/11/2003	31/12/2020
USC00467785	ROWLESBURG 1	WV	39.34	-79.67	01/01/1900	31/01/2018
USC00466867	PARSONS 3 SE	WV	39.07	-79.63	01/01/1900	01/03/2021
USC00369367	WAYNESBURG 1 E	PA	39.90	-80.17	01/05/1948	31/12/2020
USC00460094	ALBRIGHT	WV	39.49	-79.64	01/05/1953	20/04/2012
USC00465002	LAKE LYNN	WV	39.72	-79.86	01/04/1928	01/05/2021
USW00013736	MORGANTOWN MUNICIPAL AIRPORT	WV	39.65	-79.92	01/01/1900	31/12/2020
USC00367073	POINT MARION LOCK 8	PA	39.73	-79.91	01/06/1951	31/12/2020
USC00366310	NEW STANTON 1 SW	PA	40.20	-79.63	12/03/1952	31/12/2020
USW00014762	PITTSBURGH AL-LEGHENY CO AIRPORT	PA	40.36	-79.92	01/07/1935	31/12/2020
USC00363451	GRAYS LANDING	PA	39.78	-79.92	01/10/1996	31/12/2020
USC00466212	MORGANTOWN LOCK AND DAM	WV	39.62	-79.97	01/09/1921	31/12/2020

Table S42. Continued.

Station ID	Name	State	LAT(°N)	LONG(°w)	Start Date	Finish Date
USC00465626	MANNINGTON 8 WNW	WV	39.54	-80.46	12/10/1946	31/10/2019
USC00362190	DONORA 1 SW	PA	40.16	-79.87	01/08/1926	30/06/2022
USC00369050	UNIONTOWN 1 NE	PA	39.92	-79.72	01/01/1900	30/06/2018
USC00361377	CHARLEROI LOCK 4	PA	40.15	-79.90	01/01/1931	31/12/2020
USW00003802	CLARKSBURG BENEDUM AIRPORT	WV	39.30	-80.22	01/08/1939	31/12/2020
USC00461900	COOPERS ROCK STATE FOREST	WV	39.68	-79.77	01/02/1978	29/02/2012
USC00188315	SINES DEEP CREEK	MD	39.52	-79.41	01/08/1928	31/12/2020
USC00361705	CONFLUENCE 1 SW DAM	PA	39.80	-79.37	01/07/1946	31/12/2020
USC00366022	MOUNT DAVIS	PA	39.80	-79.19	01/09/1996	12/04/2011
USC00365686	MEYERSDALE 2 SSW	PA	39.78	-79.04	01/01/1963	31/10/2019
USC00362183	DONEGAL 2 NW	PA	40.13	-79.41	01/09/1943	31/12/2020
USC00365573	MCKEESPORT	PA	40.34	-79.86	18/08/1928	30/04/2020
USC00361726	CONNELLSVILLE 2 SSW	PA	40.00	-79.60	01/04/1976	31/12/2020
USC00361350	CHALK HILL 2 ENE	PA	39.85	-79.59	01/07/1977	30/11/2017
USC00366042	MOUNT PLEASANT	PA	40.22	-79.50	01/02/1981	12/03/2011
USC00368244	SOMERSET	PA	40.00	-79.08	01/09/1949	31/12/2020

Table S43. Emergent scaling factors and exponents for mean flow regimes in Monongahela River Basin, USA

Model	Statistics	Intercept	A (km2)	P_{mean} (mm)	P_{max} (mm)	BUL	VEG	AGR	AIC	NSE	RSR
	Mean	-1.62	0.99						-4.99	0.98	0.14
1	Stddev	0.10	0.03						-4.87	0.96	0.19
	P value	0.00	0.00								
	Mean	-1.71	0.93	0.24					-5.62	0.99	0.10
2	Stddev	0.09	0.03	0.07					-5.23	0.97	0.15
	P value	0.00	0.00	0.01							
	Mean	-1.86	0.94		0.18				-5.22	0.98	0.13
3	Stddev	0.19	0.04		0.10				-4.83	0.96	0.18
	P value	0.00	0.00		0.18						
	Mean	-1.16	0.88			-0.22			-5.55	0.99	0.11
4	Stddev	0.18	0.05			0.08			-5.12	0.97	0.16
	P value	0.00	0.00			0.03					
	Mean	-1.42	0.96				-0.06		-5.09	0.98	0.14
5	Stddev	0.39	0.08				0.11		-4.72	0.95	0.20
	P value	0.01	0.00				0.38				
	Mean	-1.49	0.97					-0.07	-5.39	0.99	0.12
6	Stddev	0.14	0.04					0.05	-4.77	0.94	0.20
	P value	0.00	0.00					0.03			
	Mean	-1.41	0.88	0.16		-0.13			-5.82	0.99	0.09
7	Stddev	0.20	0.05	0.07		0.07			-5.26	0.98	0.14
	P value	0.00	0.00	0.07		0.14					
	Mean	-1.62	0.93	0.19				-0.05	-5.93	0.99	0.09
8	Stddev	0.12	0.03	0.06				0.03	-5.07	0.96	0.17
	P value	0.00	0.00	0.01				0.08			
9	Mean	-1.94	0.96	0.26			0.07		-5.67	0.99	0.10

Table S43. Continued

	Stddev	0.30	0.06	0.08			0.08		-5.02	0.97	0.16
	P value	0.00	0.00	0.01			0.38				
	Mean	-1.53	0.91	0.18		-0.05		-0.04	-5.95	0.99	0.08
10	Stddev	0.20	0.05	0.07		0.10		0.05	-4.96	0.96	0.17
	P value	0.00	0.00	0.03		0.41		0.23			

Note. A, P_{mean}, P_{max}, BUL, VEG, AGR, AIC, NSE, RSR respectively refer to area, at-site mean, and maximum daily precipitation, built up land, vegetated land, agricultural land, Akaike Information Criterion coefficient, Nash-Sutcliffe Efficiency coefficient, and RMSE-observations standard deviation ratio. Calibration, and validation coefficients for AIC, NSE, RSR are represented in first and second row for each model respectively.

Table S44. Mean and Standard Deviation (in Parenthesis) of emergent scaling factors and exponents for all flow regimes (univariate model) in Monongahela River Basin, USA

Flow regimes (m3/s)	Intercept	Area (Km2)	P value (Intercept)	P value (Area)	NSE (Validation)	RSR (Validation)	AIC (Validation)
Q_{mean}	-1.62 (0.1)	0.99 (0.03)	0.00	0.00	0.96	0.19	-4.87
Q_{min}	-4.24 (1.13)	1.27 (0.36)	0.00	0.01	0.33	0.87	-0.59
Q_{1}	-4.59 (0.51)	1.5 (0.17)	0.00	0.00	0.68	0.56	-1.49
Q_{2}	-4.29 (0.42)	1.44 (0.14)	0.00	0.00	0.73	0.52	-1.74
Q_{3}	-4.07 (0.39)	1.39 (0.13)	0.00	0.00	0.76	0.49	-1.92
Q_{4}	-3.92 (0.34)	1.36 (0.11)	0.00	0.00	0.76	0.47	-2.06
Q_{5}	-3.81 (0.32)	1.34 (0.11)	0.00	0.00	0.77	0.47	-2.17
Q_{10}	-3.41 (0.28)	1.27 (0.09)	0.00	0.00	0.81	0.43	-2.47
Q_{20}	-2.88 (0.23)	1.17 (0.07)	0.00	0.00	0.86	0.36	-2.95
Q_{25}	-2.68 (0.2)	1.13 (0.06)	0.00	0.00	0.87	0.34	-3.18
Q_{30}	-2.48 (0.19)	1.1 (0.06)	0.00	0.00	0.88	0.33	-3.41
Q_{40}	-2.19 (0.15)	1.05 (0.05)	0.00	0.00	0.91	0.27	-3.85
Q_{50}	-1.99 (0.12)	1.04 (0.03)	0.00	0.00	0.93	0.24	-4.21
Q_{60}	-1.84 (0.11)	1.03 (0.03)	0.00	0.00	0.94	0.22	-4.43
Q_{70}	-1.68 (0.11)	1.02 (0.03)	0.00	0.00	0.95	0.21	-4.53
Q_{75}	-1.61 (0.11)	1.02 (0.03)	0.00	0.00	0.95	0.20	-4.59
Q_{80}	-1.53 (0.1)	1.02 (0.03)	0.00	0.00	0.96	0.19	-4.68
Q_{90}	-1.26 (0.1)	0.99 (0.03)	0.00	0.00	0.96	0.18	-4.87
Q_{95}	-1 (0.09)	0.96 (0.03)	0.00	0.00	0.96	0.17	-4.94
Q_{96}	-0.91 (0.1)	0.95 (0.03)	0.00	0.00	0.97	0.17	-4.99
Q_{97}	-0.82 (0.1)	0.93 (0.03)	0.00	0.00	0.96	0.18	-4.96
Q_{98}	-0.69 (0.1)	0.91 (0.03)	0.00	0.00	0.96	0.18	-4.93

Q_{99}	-0.52 (0.11)	0.89 (0.03)	0.00	0.00	0.95	0.21	-4.81
Q_{max}	0.26 (0.16)	0.8 (0.06)	0.23	0.00	0.81	0.41	-3.56