Watershed Prioritization in Lower Shivaliks Region of India Using Integrated Principal Component and Hierarchical Cluster Analysis Techniques: A Case of Upper Ghaggar Watershed

Nitin Chauhan[1] · Rakesh Paliwal[2] · Vipin kumar[3] · Sanjit Kumar[1] · Ritesh Kumar[1]

Abstract

The watershed prioritization of soil erosion-affected areas is an utmost requirement to formulate management and conservation practices. In this study, a geospatial framework integrated with multivariate statistical techniques such as principal component analysis and hierarchical clustering was used to prioritize the fragile ecosystem, i.e., the upper Ghaggar watershed. It is one of the least studied seasonal river watersheds in northern India, which carry substantial monsoon flows. The drainage characteristics were inferred using the interferometrically derived Sentinel 1A/1B digital elevation model (spatial resolution-13.96 m). The study area was further divided into 92 sub-watersheds (32 fourth and 60 third-order) using ArcSWAT. Twenty-seven linear, areal, and relief aspect parameters were exploited to study the watershed's hydrological, lithological, and geomorphological characteristics. The categorization and correlation of these parameters were attempted using principal component/factor analysis, which resulted in five components having an eigenvalue greater than 1. The factor analysis resulted in magnitude, relief, drainage composition, and dissection intensity factor, which accounted for percentage variance of 39.96, 27.70, 10.32, and 9.80%, respectively, which account for 87.78% of the total variance. Finally, the morphometric parameters were grouped into three priority clusters in which 29, 46, and 17 sub-watersheds fall in high, medium, and least priorities clusters, respectively. The methodology adopted in this study provides vital information for watershed characterization and prioritization, which can serve as a criterion for the decision-makers in sustainable planning and management of the resources prevalent within the watershed.

Keywords Morphometry · Principal component analysis (PCA) · Clustering · Sentinel 1 · Prioritization

Introduction

Globally, an estimated 1965 million hectares of land is subjected to one or other kind of degradation, out of which soil erosion by water and wind accounts for 1643 million hectares (Oldeman, 1991). According to the studies conducted by Rao (2000), annually, 6000 million tons of soil are lost in India. According to the Ministry of Agriculture (Soil and Water Conservation Division), 175 million hectares of vegetative land are degraded in some form or other. In India, 29.30% of the total geographical area, equivalent to 96.43 million hectares of land, is severely affected by varied land degradation (SAC, 2016). According to Singh et al. (1992), areas such as Shivalik Hills, northwestern Himalayan regions, Western Ghats, and parts of peninsular India are most severely affected by soil erosion, i.e., about 20 Mg/ha/year.

Due to increased environmental degradation with deforestation, the material carried by rivers has risen from 9.30 billion tons to 23 billion tons in a year in the last 50 years. Further, the situation becomes worse, noting that 15 billion tons a year is being contributed by Asian rivers to the ocean. Narayana and Babu (1983) observed that the Indian rivers carry and transport about 1572 million tons of soil to sea. Further, about 480 million tons of soil are

✉ Nitin Chauhan
  dr.nitinchauhan29@gmail.com

1 Haryana Space Applications Centre, Hisar, Haryana, India

2 Regional Remote Sensing Centre-West, National Remote Sensing Centre, Indian Space Research Organization, Jodhpur, India

3 School of Earth Sciences, Banasthali Vidyapith, Banasthali, Tonk, Rajasthan, India

deposited into reservoirs, reducing their storage capacity by 1–2%.

There has been a similar increase in the water scarcity problems with much severity in most parts of the world. India receives 4000 BCM (Billion Cubic meter) in form of precipitation (including snowfall and rainfall); out of this, 1869 BCM accounts for the available water. The utilizable water from available surface and groundwater accounts for 690 BCM and 490 BCM, respectively. As per the water stress index given by Falkenmark and Lindh (1976), the countries with less than 1700 cubic meters of annual per capita water availability are considered water-stressed. Similarly, when this per capita water availability goes below 1000 cubic meters, the country is declared water scarce. As per studies of the Central water commission, per capita average annual water availability in India for the year 2010 was 1588 m^3/year but by 2025 and 2050, it is going to drop to 1434 m^3/year and 1140 m^3/year, respectively (Dadhwal et al., 2012).

The watershed is considered the most viable planning unit to conserve these resources (F.A.O, 1985). The rational utilization of land and water resources, ensuring its sustainable and optimum production, and at the same time exerting minimum pressure on natural resources and the environment calls for the concept of Integrated Watershed Management. This watershed-based methodology has been logical since the land and water resources have concordant effects when developed on a watershed basis. Various techniques have evolved for watershed prioritization, out of which quantitative morphometric analysis has given promising results. Morphometry is defined as the quantitative analysis of the earth's shape, its configuration, and various landforms. It has evolved as an essential tool for identifying and prioritizing highly eroded watersheds (Nautiyal, 1994). The in-situ monitoring of soil erosion for large watersheds is very costly. Hence the geomorphometric analysis is mainly carried out using the geographical information systems approach. The regionalization of the hydrologic models is implemented effectively using these geomorphologic studies. The linear, areal, and relief characteristics have played a significant role in understanding the watershed's hydrological nature (Chow, 1964; Strahler, 1964). Due to the acceleration of watershed management programs for conservation, development, and beneficial use of natural resources such as soil and water, the demand for timely and updated information on watershed runoff and sediment yield has grown enormously in the last decade (Gajbhiye & Mishra, 2012; Meshram & Sharma, 2017; Mishra et al., 2013).

Several researchers use the conventional approach of compounding different morphometric factors for watershed prioritization and the sustainable conservation and management of the soil and water resources (Abdeta et al.,

2020; Chandniha & Kansal, 2017; Nookaratnam et al., 2005; Thakkar & Dhiman, 2007; Waiyasusri & Chotpantarat, 2020). But it has been noticed that morphometric data is inherently multivariate (Mather & Doornkamp, 1970). Therefore, the traditional statistical approaches cannot simultaneously address the similitude between the morphometric parameters. Globally, researchers have started implementing principal component analysis, factor analysis, and hierarchical clustering analysis based on multivariate techniques. They are utilizing the inherent multidimensional data of morphometry to reveal its underlying structure. Bothale et al. (1997) delineated 135 watersheds from the Bajaj Sagar dam catchment of the Mahi basin. The authors used the data reduction technique of principal component analysis (PCA) for reducing the data redundancy. The PCA resulted in 12 components, grouped lately by hierarchical cluster analysis resulting in 50 eco watersheds having similar characteristics. Gopinath et al. (2016) analyzed the watershed management planning activities for the Kuttiyadi river basin, Kerala. The basin is susceptible to flood and inundation and exhibits high runoff. The authors used a multi-criteria decision-making approach to identify the integrated effects of morphometric parameters on soil erosion. Mangan et al. (2019) utilized the correlation matrix to assess the interrelationship between the Nanganji river basin's morphometric parameters situated in Tamil Nadu. The authors implemented factor analysis and grouped the parameters into three factors, i.e., shape, magnitude, and runoff factors. Jhariya et al. (2020) prioritized the Bindra watershed in Chhattisgarh by implementing the analytical hierarchical process (AHP) based on multi-criteria decision analysis. The author delineated 16 watersheds and categorized them into very high, moderate, and low priority zones using various factors such as soil loss, land capability classification, runoff, and sediment yield in a GIS environment. Arefin et al. (2020) delineated seventeen fifth-order and three sixth-order watersheds using SRTM DEM. They utilized 16 morphometric parameters and implemented principal component analysis (PCA) for watershed prioritization. The drainage density, circularity index, elongation ratio, and bifurcation ratio were compounded for final priority. Prieto-Amparán et al. (2019) performed the quantitative multivariate geomorphic characterization of the Conchos River basin, Mexico. They delineated 31 watersheds for prioritization purposes and utilized principal component analysis (PCA), group analysis (GA), and the compounded parameter ranking methodology. The 31 watersheds were grouped into five groups based on their erosion susceptibilities potential.

All the prior studies are based on open-source digital elevation models (DEM) such as shuttle radar terrain mapping (SRTM) and advanced spaceborne thermal

emission and reflection radiometer (ASTER) onboard Terra satellite having a spatial resolution of 30 m. These existing datasets are relatively older, and morphometric phenomena are dynamic. Therefore, in this study, interferometrically generated Sentinel 1A/1B based on DEM was used to prioritize the upper Ghaggar watershed, which forms the part of one of the most fragile ecosystems in Lower Shivaliks having highly erodible soils which are lost at an alarming rate of 0 to 20,119 tones/hectare/year (Chauhan et al., 2020a). The principal component analysis (PCA) was used to reduce the parameters into five components or factors. The agglomerative hierarchical clustering technique by implementing Ward's method was used to create three priority clusters for the upper Ghaggar watershed finally.

Study Area

The study area chosen for the present study is the upper Ghaggar watershed having its confluence with the Medkhali river. It is part of India's fragile ecosystems apart from the Western and Eastern Ghats. The basin area extends from 76.86252 E and 30.61292 N to 77.21258 E and 30.90755 N and covers an area of 559.14 km^2 (Fig. 1). It covers Panchkula, Solan & Sirmaur, and Sahibzada Azad Singh Nagar districts of Haryana, Himachal Pradesh, and Punjab. The upper Ghaggar watershed has a hilly topography ranging from 249 to 1869 m. The Ghaggar river originates from the Dagshai village near Shimla, Himachal Pradesh, and is one of the most prominent ephemeral streams of the Lower Shivaliks. The river has tributaries such as the Vedic river Saraswati, Medkhali, Markanda, Tangri, and Chautang. The Ghaggar River, while flowing southwest, demarcates the boundary of Haryana and Punjab in the north. The soils of the upper Ghaggar watershed are classified into three classes, i.e., Eutric Cambisols (54.98%), Eutric Regosols (33.15%), and Eutric Fluvisols (11.88%). The study area's climate is humid subtropical and characterized by hot to sweltering summers and cool to mild winters. The average annual rainfall of the study area is 1260 mm calculated for the period ranging from 1969 to 2016. The Landuse/Landcover change analysis of the upper Ghaggar watershed conducted by Chauhan et al., (2020b) revealed that the built-up, shrubland and barren land have increased by 157.5 ha/year, 131.5 ha/year and 10.77 ha/year, respectively in the last 30 years (1985–2015) while the deciduous forest, evergreen forest and agriculture were reduced at a rate of 167.4 ha/year, 46 ha/year and 66 ha/year, respectively. The temperature of the study area varies from 4° to 37 °C. The study area also exhibits tremendously dry summers accompanied by dust storms. The geological structure of the study area is mainly composed of sandstone and conglomerate rocks.

Dataset and Software Used

All-weather, day and night capability Sentinel 1A of spatial resolution 5 m in range direction and 20 m in azimuth direction of 06.01.2016 and 13.01.2016 (Table 1) was used to generate the DEM (Digital Elevation Model). The preprocessing of the pair of single look complex (SLC) products acquired in interferometric wide (IW) mode was performed in sentinel application platform (SNAP) software (version 6.0). Sentinel images were co-registred to create a stack by utilizing the precise orbit ephemerides (POE) orbit files. One image was used as master and the other as slave, and pixel values of the slave images were moved to align with the master dataset to attain a sub-pixel accuracy to ensure the same range and azimuth was contributed by each ground object. The complex conjugate of the master image was multiplied with the slave image for the generation of the interferogram. The process of terrain observation with progressive scans (TOPS) Deburst and TOPS Merge led to the formation of the seamlessly merged single image file. The speckle filtering algorithms and Goldstein Phase Filtering (Goldstein et al., 1988) incorporated within the SNAP software were used to decrease the speckle, simultaneously maintaining radiometric information. With the implementation of Statistical-cost, Network-flow, Algorithm for Phase Unwrapping (SNAPHU), the resultant phase was generated as a result of unwrapped filtering (Chen & Zebker, 2000). The unwrapped phase was imported using the NEXT European space agency (ESA) synthetic aperture radar (SAR) toolbox (NEST) software, and DEM was prepared from it. The resultant DEM was of 13.96 m resolution against the traditional 30-90 m DEM from SRTM and was updated topographically as compared to SRTM. Its vertical accuracy was 10 m when compared with the spot heights available on the Survey of India (SOI) toposheet.

Methodology

Watershed Delineation and Drainage Analysis

The Sentinel 1 DEM was further used to generate the streams (Fig. 2) and delineate 92, i.e., sixty-third-order and thirty-two fourth-order sub-watersheds (Fig. 3), using the ArcSWAT 10.5 extension of ArcGIS 10.5. The area of individual 32 fourth-order sub-watersheds ranges from 2.749 (SW-4) to 25.862 km^2 (SW-94) while that of individual 60 third-order sub-watersheds ranges from 0.515
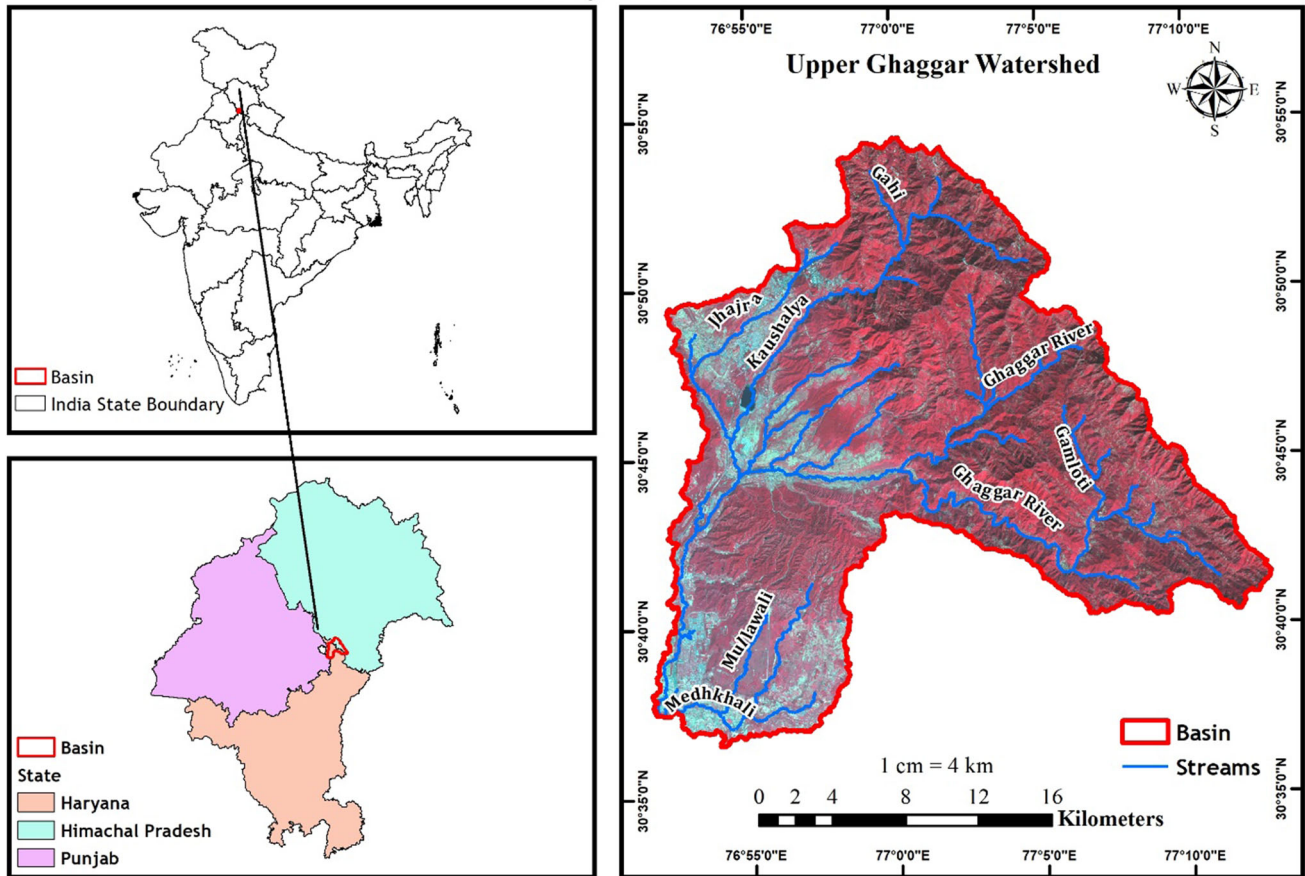
Fig. 1 Study area

Table 1 Details of the Sentinel 1 data used

Sr. no	Data/product ID	Date	Source
1	S1A_IW_SLC__1SSV_20160106T125527_20160106T125554_009374_00D90B_8F3F	06.01.2016	https://scihub.copernicus.eu/dhus
2	S1A_IW_SLC__1SSV_20160130T125526_20160130T125553_009724_00E32D_AC02	13.01.2016	

(SW-11) to 8.881 km^2 (SW-84). The ordering of the extracted streams is based on the nomenclature proposed by Strahler (1964). The linear, areal, and relief morphometric parameters analyzed during this study were calculated based on the formula shown in Table 2.

Morphometric Analysis

The linear aspects extracted for these sub-watersheds are stream order (SO_u), stream length, stream number (SN_u), mean stream length (MSL_u), Mean Stream Length Ratio ($MSLR_u$), Bifurcation ratio (R_b), and Rho Coefficient (ρ). Basin area (A), basin length (L_b), basin perimeter (P), Lemniscate's value (L_k), form factor (F_f), elongation ratio (R_e), ellipticity index (I_c), and circularity ratio (R_c) are the areal aspects extracted for these sub-watersheds. Drainage density (D_d), drainage texture (D_t), stream frequency (S_f), infiltration number (I_f), and drainage intensity (D_i) are the drainage characteristics extracted for the sub-watersheds. The relief characteristics of a basin represent the areal, volume, and altitudinal aspects of the basin landscape. The relief characteristics of morphometric analysis investigated in the study are absolute relief (R_a), dissection index (D_{is}), relative relief ratio (R_{hp}), relative relief (R_r), Ruggedness number (R_n), and slope (S).

The MSL_u acts as an essential parameter for the calculation of drainage density. The MSL_u helps in revealing the bedrock hydrological characteristics and extent of the
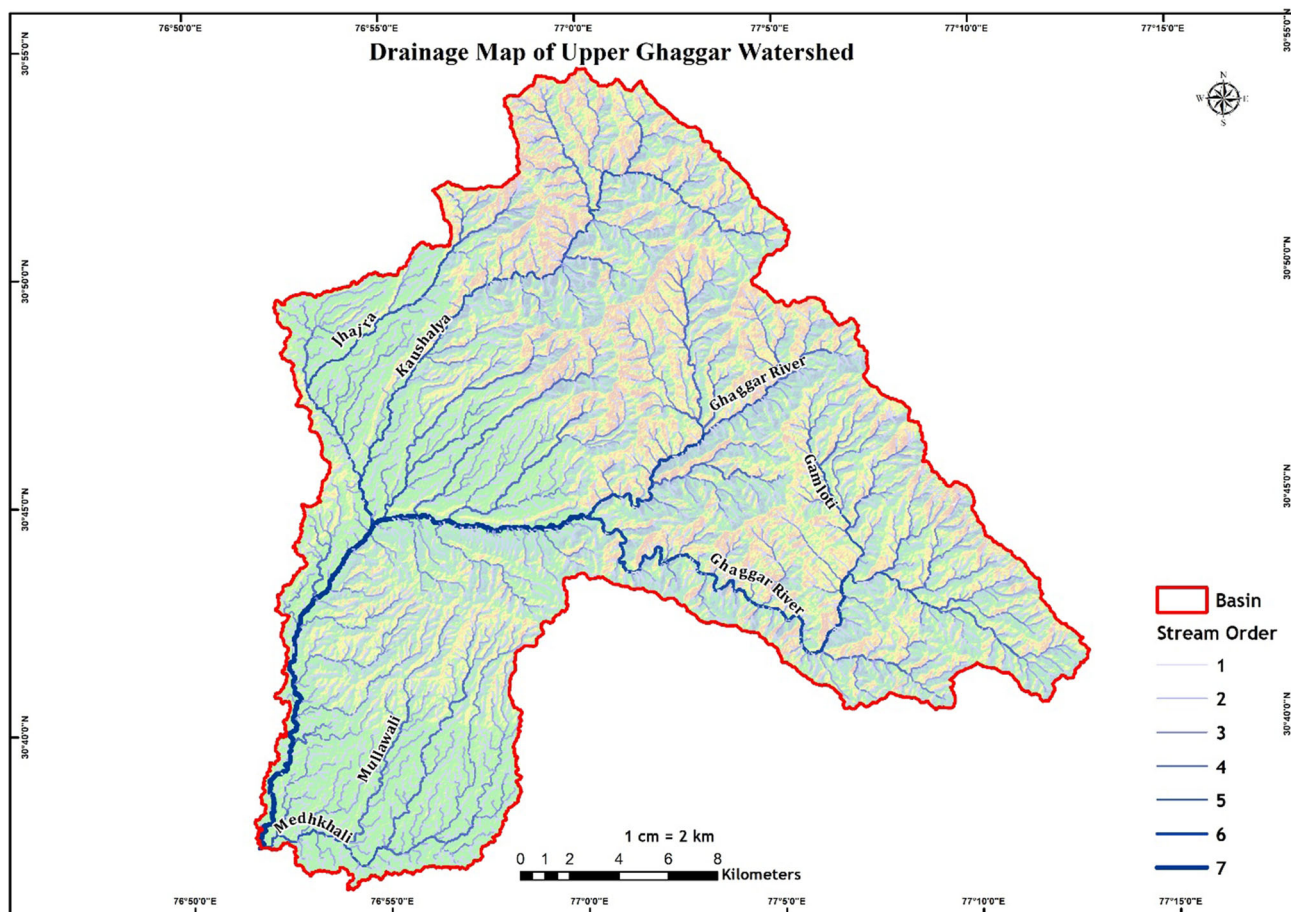
Fig. 2 Drainage map of upper Ghaggar watershed

basin. The $MSLR_u$ is an indicator of the geomorphological development stage of a basin. Horton (1945) explained that the R_b value ranges from 2 for flat or rolling basins and may go up to 3–4 for highly dissected or mountainous river basins. Strahler (1964) inferred from his study that the R_b values between 3 to 5 for the watersheds indicate that the basin's geological development is ineffective in altering the drainage patterns. The basins with higher R_b values produce a low but extended peak flow, whereas basins with lower R_b have a sharp peak flow (Strahler, 1964). The ρ helps in determining the relation of drainage composition and physiographic development of a sub-basin. The low value of ρ is indicative of low water storage during flood periods. It has a high erosion effect, while the higher value is indicative of higher hydrologic storage during floods and thus reduces erosion effects at peak discharges. The L_k values drastically control the shape of the drainage basin. The L_k values are one for circular basins, and as its value increases, the shape of the drainage basins becomes more and more elongated. An elongated basin is less effective in runoff discharge than a circular basin (Singh & Singh, 1997). The F_f is an indicative morphometric parameter of

the flood-regime of the stream in the case of long and elongated drainage basins (Horton, 1932). The value of F_f varies from 0 for highly elongated basins to unity, i.e., 1 for perfectly circular shaped basins. R_e acts as an imperative index for basin shape analysis, and the areas having higher values possess high infiltration capacity and low surface runoff. The I_e provides a close relationship between morphometry and hydrology (Stoddart, 1965). Lower I_e values indicate a quick runoff draining basin because of which the stream channels might swell or overflow, resulting in downstream flooding in case of heavy rainfall. R_c defines the circularity of the basin and is a dimensionless entity. R_c values range from 0 to 1 with zero value given to basins having a highly elongated shape to one value given to circular basins.

D_d acts as a permeability indicator of the drainage basin surface. High D_d is pertinent to the regions having impervious subsurface with weak structure resulting in more runoff. In contrast, low D_d is associated with highly resistant subsurface covered by vegetation and areas with low relief, resulting in lower runoff (Prasad et al., 2008). Singh (1976) defined the drainage texture based on the
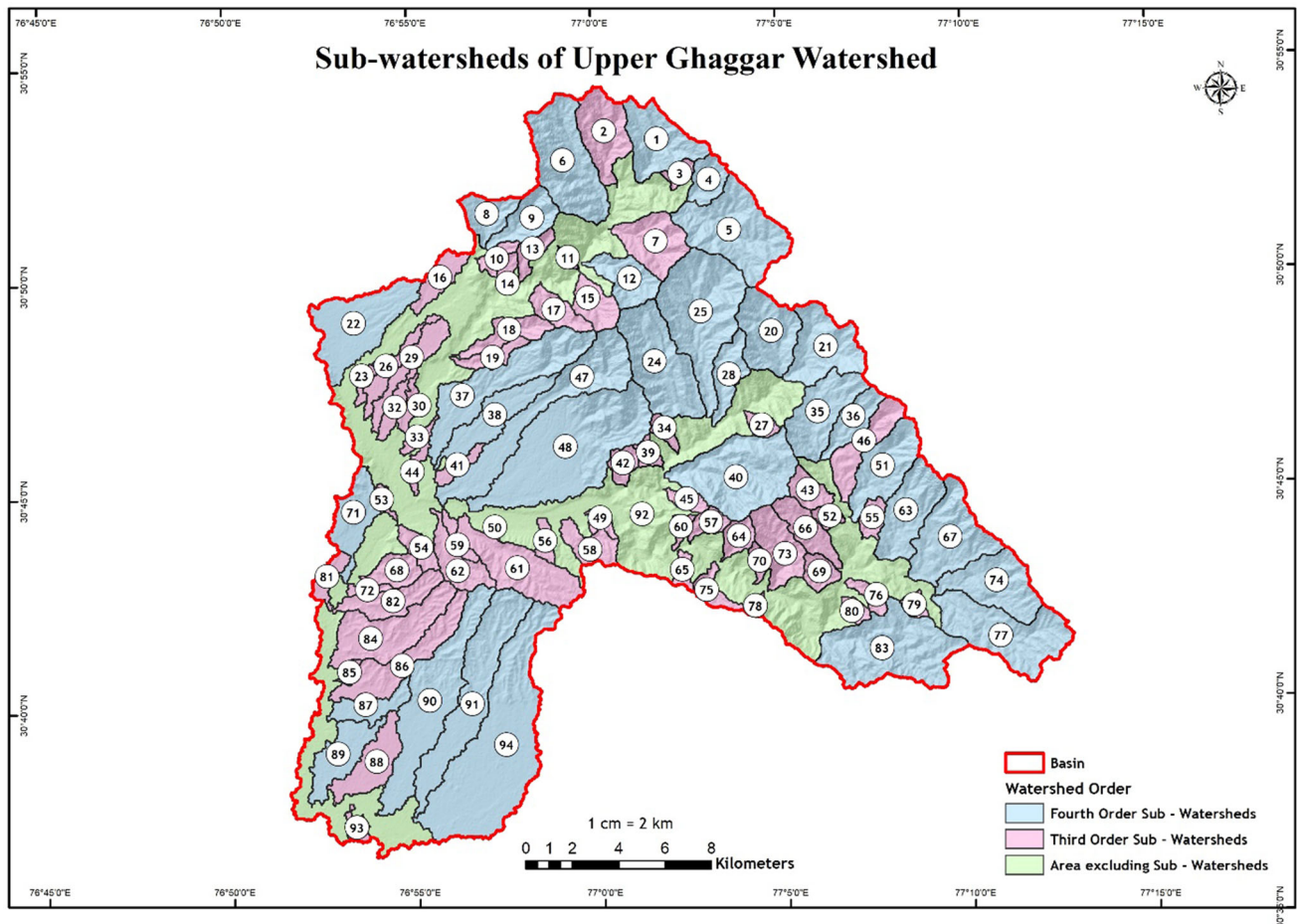
Fig. 3 Ninety-two sub-watersheds of upper Ghaggar watershed

relative spacing of the streams. D_t is an indicator of the infiltration capacity, rock permeability, and relief aspect. The fine texture is prevalent in areas having low permeability and less resistance to erosion, and in the case of coarse texture areas, it is vice-versa. S_f shows a positive correlation with the drainage density and is, therefore, a similar indicator of infiltration capacity, rock permeability, and basin relief. High S_f values in hilly regions are associated with steep slopes and more significant rainfall, while in plain areas, lesser surface flow and low permeability result in low S_f value. Faniran (1968) defined the I_f of a drainage basin as the product of D_d and S_f. The higher values of I_f contribute to lower infiltration rates and result in higher runoff. D_i was defined by Faniran (1968) as the ratio of the S_f to the D_d. It is an indicator of the extent to which the denudational agents depress the surface under drainage density and stream frequency effect.

R_a defines the maximum elevation of a basin area. This parameter helps in determining the rate of erosion concerning the recent summit or hilltops of a basin. Higher R_a is associated with high erosional activities. Smith (1935) coined the term R_r to the highest and lowest altitude points

of a particular area. R_r plays a vital role in calculating average slope, dissection determination, and assessing the terrain development stages. Melton (1958) gave the R_{hp} as the ratio of the R_r and P of the area. R_{hp} ratio acts as an indicator of the relative velocity of the vertical tectonic movements. The lower values of R_{hp} pertain to the less resistive rocks. D_{is} is the ratio of R_r to R_a of an area. It acts as a vital parameter for developing an understanding of the degree of dissection and evolution of landform development stages in any given physiographic region.

The values of D_{is} range from 0, which implies a theoretical value as there is no region in nature that is passive to erosion to 1. R_n measures the structural complexity of the terrain. It is a dimensionless property calculated as a product of R_r and D_d of a given basin area having the same units. Patton and Baker (1976) discussed that the regions having higher ruggedness numbers accompanied with fine drainage texture and minimalistic length of overland flow on steep slopes have the expectancy of potential flash flooding. These morphometric parameters combination may lead to higher flood peaks for an area having a low ruggedness number even for the equivalent of rainfall

Table 2 Methods used to calculate the morphometric parameter

Sr. no	Parameter	Formula	References
1	Stream order	SOu = Hierarchical rank	Strahler (1957)
2	Stream number	$SN_u = SN_1 + SN_2 + \cdots SN_n$	Horton (1945)
		Where SN_1 = Stream number of first order and so on and SN_u = Total no. of stream segments of order 'u'	
3	Stream length (SL_u)	$SL_u = SL_1 + SL_2 \ldots \ldots SL_n$	Strahler (1964)
		Where SL_1 = Stream length of first order and so on and SL_u = Total stream length of order 1 to n	
4	Mean stream length (MSL_u)	$MSL_u = \dfrac{\sum_{i=1}^{N} SL_u}{SN_u}$	Strahler (1964)
		Where SL_u = Total stream length of order 'u' and SN_u = Total no. of stream segments of order 'u'	
5	Mean stream length ratio ($MSLR_u$)	$MSLR_u = \dfrac{MSL_u}{MSL_{u-1}}$	Horton (1945)
		Where MSL_u = Mean stream length of a given order and	
		MSL_{u-1} = Mean stream length of next lower order	
6	Weighted mean stream length ratio ($WMSLR_u$)	$WMSLR_u = \dfrac{\sum_{i=1}^{n} MSLR_u \times SLIR_u}{\sum_{i=1}^{n} SLIR_u}$	Strahler (1952)
		Where $MSLRu$ = Mean Stream Length ratio of a given order	
		$SLIRu$ = Stream length involved in the ratio	
7	Bifurcation ratio (R_b)	$R_b = \dfrac{SN_u}{SN_{u+1}}$	Strahler (1964)
		Where SN_u = Stream number of given order	
		SN_{u+1} = Stream number of next higher order	
8	Weighted mean bifurcation ratio (WMR_b)	$WMR_b = \dfrac{\sum_{i=1}^{n} R_b \times SNR_b}{\sum_{i=1}^{n} SNR_b}$	Strahler (1953)
		Where R_b = Bifurcation ratio, SNR_b = Total number of Streams involved in the ratio	
9	Rho coefficient (ρ)	$\rho = \dfrac{MSL_{R_u}}{R_b}$	Horton (1945)
		Where MSL_{R_u} is stream length ratio,	
		R_b is bifurcation ratio	
10	Basin area (A)	ArcGIS/DEM	Schumm (1956)
11	Perimeter (P)	ArcGIS/DEM	
12	Basin length (L_b)	$L_b = 1.312 A^{0.568}$	Schumm (1956), Gardiner (1975), Nookaratnam et al., (2005) and
		Where A = Basin Area	
13	Lemniscate's value (L_k)	$L_k = \dfrac{L_b^2 \pi}{4A}$	Chorley et al., (1957)
		Where L_b = Basin Length, A = Basin Area	
14	Form factor (F_f)	$F_f = \dfrac{A}{L_b^2}$	Horton (1932)
		Where L_b = Basin Length, A = Basin Area	
15	Elongation ratio (R_e)	$R_e = \dfrac{1}{L_b} \times \sqrt{\dfrac{4A}{\pi}}$	Schumm (1956)
		Where L_b = Basin Length, A = Basin Area	
16	Ellipticity index (I_e)	$I_e = \dfrac{\pi VL^2}{4A}$	Stoddart (1965)
		Where VL = Valley Length, A = Basin Area	
17	Circularity ratio (R_c)	$R_c = \dfrac{4\pi A}{P^2}$	Miller (1953)
		Where P = Basin Perimeter (km), A = Basin Area (km^2)	
18	Drainage density (D_d)	$D_d = \dfrac{\sum SL_u}{A}$	Horton (1932)
		Where $\sum SL_u$ = Total Length of all stream segments, A = Total area of the basin	

Table 2 (continued)

Sr. no	Parameter	Formula	References
19	Drainage texture (D_t)	$D_t = \dfrac{1}{(t + P)/2}$ $t = \dfrac{(t_1 + t_2)/2}{\sqrt{2}}$ $P = \dfrac{P_1 + P_2 + P_3 + P_4}{4}$ Where t_1 & t_2 = number of intersections between the drainage network and grid diagonal P_1 to P_4 = number of intersections between the drainage network and grid edges	Horton (1945) and Singh (1976)
20	Stream frequency (S_f)	$S_f = \dfrac{\sum_{i=1}^{K} SN_u}{A}$ Where $\sum SN_u$ = Total number of all stream order, A = Total area of the basin	Horton (1945)
21	Drainage intensity (D_i)	$D_i = \dfrac{S_f}{D_d}$ Where S_f = Total number of all stream order, D_d = Drainage density	Faniran (1968)
22	Infiltration number (I_f)	$I_f = D_d \times S_f$ Where S_f = Total number of all stream order, D_d = Drainage density	Faniran (1968)
23	Relative relief (R_r)	R_r = Max. Elevation − Min. Elevation	Smith (1935)
24	Relative relief ratio (R_{hp})	$R_{hp} = 100 \times \dfrac{R_r}{\text{Perimeter of the basin}}$ Where R_r = Ralative Relief	Melton (1958)
25	Dissection index (D_{is})	$D_{is} = \dfrac{R_r}{R_a}$ Where R_r = Relative relief, R_a = Absolute relief	Nir (1957)
26	Ruggedness number (R_n)	$R_n = R_r \times D_d$ Where R_r = Relative relief, D_d = Drainage density	Strahler (1964)
27	Slope	ArcGIS/DEM	

events. The slope is an invariable relief morphometric parameter related to the infiltration capacity and runoff of an area. It is inversely proportional to the infiltration capacity. The mean values of linear, areal, and relief morphometric parameters for 92 sub-watersheds are given in Online Resource 1 (ESM_1.docx).

Principal Component Analysis of the Morphometric Parameters

The interrelationship between the different morphometric parameters was analyzed using a correlation matrix. To further strengthen the correlation matrix analysis results and give a clearer insight, a data reduction or factor analysis of the above-given parameters was done using the principal component analysis (PCA) in statistical package for the social sciences (SPSS) software version 25.

Principal component analysis (PCA) is a multivariate data dimensionality reduction technique, which reduces the data into fewer factors but represents the entire dataset. The factors having an eigenvalue larger than one were used to maintain the significance of the contributing factor, and as a result, only five components were used. A normalized varimax rotation was applied on the resultant matrix to compute the factor loadings between the original geomorphic parameters and the final components/factors. The factor loading governs the correlation's strength, i.e., a higher factor loading value means stronger correlation and vice versa. All the score factors of the 92 sub-watersheds were divided into four factors, i.e., magnitude, relief, drainage composition, and dissection intensity factors depending on the morphometric aspects, which define 87.8% of the total variance.

Watershed Prioritization Using Hierarchical Clustering Analysis

In the study, agglomerative Hierarchical Cluster Analysis using Ward's method (Ward, 1963) was applied for identifying the natural grouping within a given dataset for watershed prioritization. The more related data come together to form a homogeneous group while the dissimilar data become part of some other contrasting group. Ward's method (Ward, 1963) is also called the minimum variance method because of its ability to create even-sized compact clusters but is computationally intensive. It offers the advantage of separating the cluster if the variance among the cluster is higher than the defined threshold. The squared Euclidean distance was used as the distance parameter to assess the distance between the respective cluster.

Results and Discussion

Morphometric Analysis

The morphometric analysis of the upper Ghaggar watershed reveals that it is a seventh-order watershed. It consists of 3483 first, 737 s, 160 third, 32 fourth, 9 fifths, 3 sixth, and 1 seventh order stream, suggesting that these stream channels generate a significant surface runoff and sediment yield (Bhat et al., 2019; Obeidat et al., 2021). The stream length of the basin is higher for its first order and decreases with an increase in stream order. Any deviation of this trend is due to inherent variations of topography and slope. The upper Ghaggar watershed's stream lengths for its 1st, 2nd, 3rd, 4th, 5th, 6th, and 7th orders are shown in Table 3. The Upper Ghaggar watershed's mean stream length for its 1st, 2nd, 3rd, 4th, 5th, 6th, and 7th orders are shown in Table 4. Rai et al. (2017) indicated in their study that low values of MSL_u are mostly associated with the mountainous environment. The low values of MSLu represent the high erosion potential and younger geomorphological development of the landscape. The general trend of MSL_u increases as the stream order increases, but any abnormalities in this suggest changes in the underlying slope and geology, which affects the flow characteristics. The $MSLR_u$ of the upper Ghaggar watershed displays an increasing trend from smaller order to higher order indicating their matured geomorphological development stage as against the sub-watersheds displaying abrupt changes between the orders indicating the late youth stage of geomorphological development. The $MSLR_u$ of the upper Ghaggar watershed for 1st to 2nd, 2nd to 3rd, 3rd to 4th, 4th to 5th, 5th to 6th, and 6th to 7th, respectively are shown in Table 5. Mean R_b values for the upper Ghaggar watershed for 1st to 2nd, 2nd to 3rd, 3rd to 4th, 4th to 5th, 5th to 6th, and 6th to 7th, respectively are shown in Table 6. The inconsistent values across the orders are due to the strong geological and lithological impact over the basin area. The Rho coefficient value of the upper Ghaggar watershed is 0.56, indicating less water storage and high erosion during the flood.

The L_k, F_f, R_e, and I_e values of the upper Ghaggar watershed are 3.20, 0.25, 0.56, and 2.27, respectively, displaying its elongated shape. The higher values of L_k suggest a potentially high erosive nature of the basin. The low F_f represents a flatter peak of flow for a longer duration, which facilitates groundwater percolation. The elongated basins are less effective in runoff discharges because of their high elevation and steep gradient (Rai et al., 2018). The D_d of the upper Ghaggar watershed is 3.15 (km/km^2), having a medium drainage density. The upper Ghaggar watershed is having a coarser D_t (0.62), higher S_f (11.59), poor D_i (3.94), and higher I_f (39.04), resulting in a higher runoff.

The average R_a and R_r of the upper Ghaggar watershed are 902.48 m and 221.56 m, respectively. The average R_{hp} of the upper Ghaggar watershed is 0.11 indicator of less resistive rocks resulting in a higher runoff. The average D_{is} of the upper Ghaggar watershed is 0.22. A moderately high R_n (0.62) is shown by the upper Ghaggar watershed with a higher potential for flash flooding. The average S value of the upper Ghaggar watershed is 15.84, which indicates a steep slope representing high erosion operable within the basin.

Correlation Analysis of Morphometric Parameters

The interrelationship between the 27 morphometric parameters for the 32 forth order sub-watersheds and 60

Table 3 Stream length of upper Ghaggar watershed

Watershed	Stream order							Total length of streams in upper Ghaggar watershed
	1	2	3	4	5	6	7	
Ghaggar	879.98	476.14	249.23	122.5	64.23	32.39	27.64	1852.11

Table 4 Mean stream length of upper Ghaggar watershed

Watershed	Stream order							Total mean length of streams in upper ghaggar watershed
	1	2	3	4	5	6	7	
Ghaggar	0.26	0.65	1.56	3.83	7.14	10.80	27.64	**51.85**

Table 5 Stream length ratio of upper Ghaggar watershed

Watershed	Stream length ratio						Average Stream length ratio of upper Ghaggar watershed
	2/1	3/2	4/3	5/4	6/5	7/6	
Ghaggar	2.56	2.41	2.46	1.86	1.51	2.56	2.23

Table 6 Mean bifurcation ratio of upper Ghaggar watershed

Watershed	Bifurcation ratio						Average bifurcation ratio
	1/2	2/3	3/4	4/5	5/6	6/7	
Ghaggar	4.73	4.61	5.00	3.56	3.00	3.00	3.98

third-order sub-watersheds were analyzed to assess the mutual interdependence prevalent among the different parameters shown in Table 7. It is inferred from the correlation analysis that most of the morphometric parameters are positively correlated to each other. Further, there is no single dependent variable, but all the variables influence each other and are interconnected. The correlation matrix depicts that there exists strong (0.8–0.9), Good (0.7–0.8), moderate (0.5–0.7), and also negative correlation. Basin Area (A), being the critical morphometric parameter responsible for controlling the peak flow and the amount of sediment being eroded, strongly correlates with SN_u, SL_u, MSL_u, R_b, P, L_b, and L_k. It is evident that as the area increases, so does the SN_u and SL_u. Basin area is strongly positively correlated to L_b, so an increase in the area leads to an increase in L_b.

Similarly, the area also defines the watershed's shape since it is inversely proportional to L_k, so watersheds having lower L_k values are more circular and therefore have more area. The L_b shows a strong correlation with SN_u, SL_u, MSL_u, R_b, A, and P. Strahler (1964) further indicated that basins with higher R_b values are more elongated and, therefore, produced a low but extended peak flow. In contrast, basins with lower R_b are more circular and produce a sharp peak flow. The F_f is strongly correlated with the R_e. The value of F_f varies from 0 for highly elongated basins to unity, i.e., 1 for perfectly circular shaped basins. It is shown that the sub-watersheds are having slightly elongated shapes having low F_f with a flatter peak of flow for longer duration resulting in

groundwater percolation. The R_c shows a good correlation with the relief parameters such as S, while it shows a moderate correlation with R_a, R_r, R_{hp}, and R_n. R_c attributes to the high to moderate relief and structurally controlled drainage system. The D_d shows a strong correlation with I_f. Although the inter-correlation matrix was helpful, it induced a larger number of components into fewer components based upon their importance (Siddiqui et al., 2020). Hence, the inter-correlation matrix is subjected to Principal component analysis (PCA) for categorizing the 27 morphometric parameters into 5 principal components.

Multivariate Analysis Using the Principal Component Approach

The PCA resulted in 5 significant components having an eigenvalue greater than one, as shown in Fig. 4. The eigenvalues for PC1, PC2, PC3, PC4, and PC5 are 10.79, 7.48, 2.79, 1.57 and 1.07, respectively. They account for percentage variance of 39.96, 27.70, 10.32, 5.82, and 3.98, respectively, which account for 87.78% of the total variance explained by 27 morphometric parameters. The component-wise eigenvalues, variance percentage, and cumulative variance percentages are shown in Table 8. The most correlated and loaded variables in a particular component displayed in bold are shown in the varimax rotated component matrix Table 9. The PC_1 shows the highest factor loading for positively SN_u, SL_u, MSL_u, A, P, L_b, and L_k, contributing to 39.96% of the total variance and designated as the sub-watershed magnitude factor

Table 7 Pearson Correlation Matrix for the 25 parameters for sub-watersheds of upper Ghaggar watershed

	N_u	SL_u	MSL_u	$MSLR_u$	$WMSLR_u$	R_b	WMR_b	ρ	A	P	L_b	L_k	F_f	R_e	I_e	R_c	D_d	S_f	D_t	Di	I_f	R_a	R_r	R_{hp}	D_{is}	R_n	S
N_u	1.00																										
SL_u	0.97	1.00																									
MSL_u	0.85	0.88	1.00																								
$MSLR_u$	0.25	0.24	0.51	1.00																							
$WMSLR_u$	0.23	0.24	0.48	0.97	1.00																						
R_b	0.51	0.49	0.63	0.74	0.74	1.00																					
WMR_b	0.40	0.40	0.53	0.55	0.58	0.87	1.00																				
ρ	-0.12	-0.11	0.09	0.66	0.62	0.03	-0.07	1.00																			
A	0.99	0.98	0.88	0.26	0.25	0.53	0.43	-0.12	1.00																		
P	0.90	0.93	0.96	0.39	0.39	0.58	0.50	0.00	0.92	1.00																	
L_b	0.97	0.95	0.90	0.32	0.31	0.60	0.51	-0.12	0.98	0.94	1.00																
L_k	0.91	0.88	0.88	0.37	0.37	0.65	0.58	-0.10	0.92	0.92	0.98	1.00															
F_f	-0.86	-0.82	-0.85	-0.40	-0.40	-0.67	-0.61	0.08	-0.87	-0.89	-0.95	-0.99	1.00														
R_e	-0.87	-0.83	-0.86	-0.40	-0.40	-0.67	-0.62	0.09	-0.87	-0.90	-0.95	-0.99	1.00	1.00													
I_e	0.27	0.39	0.64	0.48	0.50	0.35	0.30	0.33	0.31	0.62	0.34	0.37	-0.37	-0.38	1.00												
R_c	-0.16	-0.27	-0.45	-0.36	-0.40	-0.20	-0.19	-0.34	-0.19	-0.48	-0.21	-0.23	0.23	0.23	-0.87	1.00											
D_d	-0.01	0.11	0.10	-0.07	-0.05	-0.15	-0.21	0.04	0.00	0.12	-0.06	-0.13	0.17	0.16	0.44	-0.50	1.00										
S_f	-0.10	-0.12	-0.18	-0.10	-0.16	-0.28	-0.36	0.12	-0.16	-0.18	-0.21	-0.28	0.31	0.31	-0.07	0.03	0.33	1.00									
D_t	0.09	0.00	-0.06	-0.05	-0.05	0.10	0.17	-0.13	0.07	-0.01	0.13	0.19	-0.23	-0.22	-0.35	0.25	-0.49	-0.19	1.00								
Di	0.03	-0.09	-0.16	-0.08	-0.10	-0.05	-0.02	-0.06	-0.01	-0.17	0.02	0.05	-0.07	-0.07	-0.54	0.59	-0.71	0.09	0.51	1.00							
I_f	-0.10	-0.02	-0.05	-0.13	-0.12	-0.24	-0.32	0.02	-0.11	-0.03	-0.18	-0.25	0.29	0.29	0.28	-0.36	0.83	0.67	-0.41	-0.48	1.00						
R_a	0.12	-0.03	-0.10	-0.13	-0.20	0.04	0.11	-0.20	0.08	-0.09	0.13	0.18	-0.20	-0.19	-0.54	0.61	-0.71	-0.04	0.40	0.74	-0.59	1.00					
R_r	-0.01	-0.14	-0.19	-0.19	-0.26	-0.06	-0.02	-0.20	-0.04	-0.22	-0.01	0.01	-0.02	-0.03	-0.58	0.67	-0.70	-0.04	0.34	0.72	-0.59	0.92	1.00				
R_{hp}	-0.42	-0.46	-0.55	-0.36	-0.44	-0.40	-0.36	-0.15	-0.44	-0.59	-0.49	-0.53	0.54	0.53	-0.63	0.70	-0.39	0.15	0.07	0.48	-0.27	0.52	0.73	1.00			
D_{is}	-0.19	-0.26	-0.27	-0.16	-0.23	-0.13	-0.15	-0.14	-0.20	-0.35	-0.20	-0.20	0.20	0.19	-0.50	0.59	-0.61	-0.07	0.23	0.56	-0.49	0.61	0.83	0.77	1.00		
R_n	-0.06	-0.17	-0.23	-0.22	-0.30	-0.10	-0.07	-0.23	-0.09	-0.26	-0.07	-0.06	0.05	0.05	-0.57	0.67	-0.58	0.03	0.22	0.62	-0.48	0.86	0.97	0.77	0.83	1.00	
S	0.00	-0.14	-0.17	-0.12	-0.20	0.00	0.03	-0.19	-0.02	-0.21	0.01	0.03	-0.05	-0.05	-0.60	0.71	-0.79	-0.10	0.36	0.75	-0.66	0.89	0.95	0.68	0.85	0.90	1.00

	Strong	0.8 - 0.9		Good	0.7 - 0.8		Moderate	0.6 - 0.7
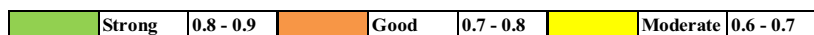
Fig. 4 Scree plot illustrating the 5-component solution resulting from PCA

Scree Plot

(Eigenvalue vs. Component Number: 10.79, 7.48, 2.79, 1.57, 1.07, ...)
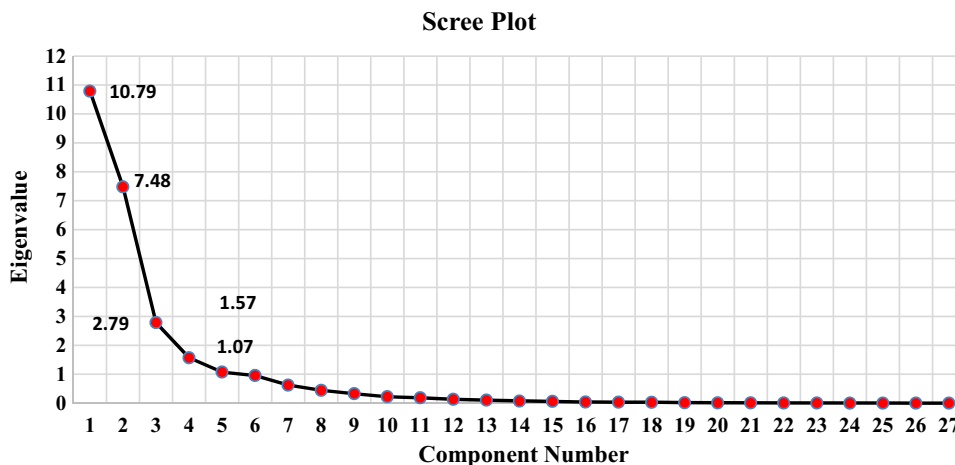
corresponding to the watershed dimensions. The PC_2 shows the highest factor loading for positively correlated R_c, D_i, R_a, R_r, R_{hp}, D_{is}, R_n, and S, contributing to 27.70% of the total variance and designated as the Relief factor corresponding to the basin relief and slope steepness.

This factor illustrates that the sub-watersheds' hydrological response behavior and higher value or steeper slopes are more prone to runoff and soil losses. The PC_3 shows the highest factor loading for positively correlated $MSLR_u$, and ρ contributes to 10.32% of the total variance and can be designated as the Drainage composition factor. This factor illustrates the Channel storage capacity, which controls flood crest intensities downstream and is necessary for flood routing and control (Horton, 1945). The PC_4 and PC_5 show the highest factor loading for positively correlated S_f and D_t, contributing to 5.82% and 3.98%, respectively of the total variance and designated as the Dissection Intensity factor. The factor scores for the 92 (32

Table 8 Component wise explanation of total variance

Component	Initial eigenvalues			Extraction sums of squared loadings			Rotation sums of squared loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	10.80	40.00	40.00	10.80	40.00	40.00	9.30	34.40	34.40
2	7.50	27.70	67.70	7.50	27.70	67.70	7.40	27.40	61.80
3	2.80	10.30	78.00	2.80	10.30	78.00	3.30	12.20	74.00
4	1.60	5.80	83.80	1.60	5.80	83.80	2.20	8.00	82.00
5	1.10	4.00	87.80	1.10	4.00	87.80	1.60	5.80	87.80
6	1.00	3.50	91.30						
7	0.60	2.30	93.60						
8	0.40	1.60	95.30						
9	0.30	1.20	96.50						
10	0.20	0.80	97.30						
11	0.20	0.70	98.00						
12	0.10	0.50	98.50						
13	0.10	0.40	98.90						
14	0.10	0.30	99.20						
15	0.10	0.20	99.40						
16	0.00	0.10	99.50						
17	0.00	0.10	99.70						
18	0.00	0.10	99.80						
19	0.00	0.10	99.80						
20	0.00	0.00	99.90						
21	0.00	0.00	99.90						
22	0.00	0.00	100.00						
23	0.00	0.00	100.00						
24	0.00	0.00	100.00						
25	0.00	0.00	100.00						
26	0.00	0.00	100.00						
27	0.00	0.00	100.00						

Fourth + 60 Third Order) sub-watersheds of the upper Ghaggar watershed are shown in Table 10. The bold values show the highest among the magnitude, relief, drainage composition, and Dissection Intensity factors for each of the 92 (32 Fourth + 60 Third Order) sub-watersheds. The dominating factor scores prevalent within the 92 sub-watersheds are shown in Fig. 5.

Prioritization Using Hierarchical Clustering Analysis

The clustering analysis resulted in three prominent clusters and is represented as Dendrogram (Fig. 6). In the Dendrogram, the sub-watersheds are plotted on the vertical axis, and the relative cluster difference represented by linkage distances are plotted on the horizontal axis. The Cluster wise mean values of the prominent linear, sub-

watershed geometry characteristics, drainage, and relief characteristics, and their respective ranks are mentioned in the brackets are listed in Table 11. The Upper Ghaggar watershed and its sub-watersheds have dendritic drainage patterns, and these kinds of patterns are prominent where the underlying rock structure does not firmly control the stream channels, and channels follow the slope of the terrain. There is a high proportion of first-order streams in Cluster 3 sub-watersheds and indicating the structural weakness prevalent in the sub-watersheds associated with these clusters.

Cluster Id 3 is having the highest values of R_b, which is indicative of highly dissected sub-watersheds. The higher R_b values correspond to sub-watersheds significantly influenced by drainage network, whereas its low values are associated with sub-watersheds having minimal disturbing drainage patterns structurally less disturbed. The ρ is

Table 9 Varimax rotated component matrix for 27 variables distributed in five principal components

	Component				
	1	2	3	4	5
N_u	**0.97**	0.00	− 0.01	0.06	0.03
SL_u	**0.96**	− 0.12	− 0.03	0.04	− 0.07
MSL_u	**0.90**	− 0.17	0.27	− 0.07	− 0.17
$MSLR_u$	0.26	− 0.09	**0.93**	− 0.12	− 0.04
$WMSLR_u$	0.25	− 0.18	**0.91**	− 0.18	0.01
R_b	0.56	− 0.02	0.53	− 0.39	0.01
WMR_b	0.48	− 0.02	0.40	− 0.53	0.08
ρ	− 0.17	− 0.13	**0.80**	0.24	− 0.01
A	**0.98**	− 0.03	− 0.01	0.00	− 0.01
P	**0.94**	− 0.22	0.14	− 0.06	− 0.10
L_b	**0.99**	− 0.01	0.05	− 0.08	0.05
L_k	**0.96**	0.00	0.12	− 0.17	0.11
F_f	− 0.93	0.00	− 0.16	0.22	− 0.16
R_e	− 0.93	0.00	− 0.15	0.23	− 0.14
I_e	0.36	− 0.58	0.40	− 0.07	− 0.38
R_c	− 0.21	**0.72**	− 0.31	0.03	0.21
D_d	− 0.01	− 0.73	− 0.14	0.37	− 0.39
S_f	− 0.12	0.00	0.04	**0.91**	− 0.03
D_t	0.09	0.25	− 0.04	− 0.16	**0.83**
D_i	0.03	**0.74**	0.05	0.15	0.50
I_f	− 0.11	− 0.59	− 0.11	0.65	− 0.22
R_a	0.15	**0.88**	− 0.07	− 0.04	0.23
R_r	0.01	**0.97**	− 0.09	− 0.04	0.05
R_{hp}	− 0.46	**0.76**	− 0.19	0.12	− 0.16
D_{is}	− 0.19	**0.87**	− 0.05	− 0.08	− 0.12
R_n	− 0.04	**0.95**	− 0.13	0.02	− 0.09
S	0.01	**0.97**	− 0.04	− 0.11	0.12

dependent on hydrologic, geologic, and physiographic factors, which ultimately determine the relation of drainage composition and physiographic development of a sub-watershed. The low value of ρ indicates low water storage during flood periods and has a high erosion effect, while the higher value indicates higher hydrologic storage during floods and thus reduces erosion effects at peak discharges. Hence, the lowest ρ values are in sub-watersheds associated with Cluster 1.

The shape of sub-watersheds is also a prominent factor in governing the prioritization since it controls the water flow rate to the main channel. In hierarchical cluster analysis, L_k, F_f, R_e, I_e, and R_c were used. The sub-watersheds belonging to cluster 3 have low F_f, R_e, I_e, and R_c are shown in Table 11, which implies that they have an elongated shape. According to Singh and Singh (1997), an elongated watershed is less effective in runoff discharge than the circular watershed. These elongated sub-watersheds have a flatter peak of flow for a longer duration, which results in high infiltration and low surface runoff. The sub-watersheds belonging to Cluster 1 have relatively high F_f, R_e, I_e, and R_c in Table 11, indicating its circular shape, which further implies high peak flows and shorter lag time than elongated sub-watersheds. Circular sub-watersheds are at higher risk of erosion and simultaneously higher sediment load since the runoff is contributed by the entire area, putting these sub-watersheds at higher risk of flash floods.

The morphometric parameters such as D_d, S_f, D_t, D_i, and I_f indicate the landscape's extent being dissected by the stream network. These help in the prediction of sub-watershed surface runoff, its sediment yield, and terrain dissection magnitude. The drainage texture of the 3 clusters varies from moderate to coarser. However, according to the area's geology, the entire watershed's 60% area is under Shale, Sandstone, and Limestone characterizing its low permeability and its erosion-prone nature. High values of D_d, S_f, and I_f are present in Cluster 3, whereas low values prevail in Cluster 1, as shown in Table 11. The high values of Cluster 3 are characterized by steep slopes, high relief, weak and impervious subsurface resulting in high surface runoff. The low values of Cluster 1 are characterized by a gentle slope, low relief, resistance and permeable subsurface, resulting in less surface runoff and providing excellent groundwater recharge sites.

The relief characteristics define the terrain's erosional properties, landform, development of drainage network, and overland flow. In the study, morphometric parameters such as R_a, R_r, R_{hp}, D_{is}, R_n, and S were used to reveal the watershed's erosion potential and runoff. The R_a and R_r of the sub-watersheds coming under Cluster 1 are 1321.11 m indicating its high erosional capacity to move water and sediments down the slope. The R_{hp} ranges from 0.37 to 3.22, with sub-watersheds coming under Cluster 3 having low values due to its low slopes and resistant bedrock, whereas Cluster 2 and Cluster 1 have relatively higher values (Table 11). Higher values of R_{hp} characterize steep slopes, high erosion intensity, and higher sediment

Table 10 Factor scores for 92 (32 fourth and 60 third-order) sub-watersheds of upper Ghaggar watershed

Sub-watersheds	Factor 1 Magnitude factor	Factor 2 Relief factor	Factor 3 Drainage composition factor	Factor 4 Dissection intensity factor	Factor 5
1	0.67	0.39	− 0.13	0.43	**2.85**
2	0.40	0.68	**1.16**	− 1.82	0.88
3	− 0.87	**1.34**	0.39	0.66	− 0.40
4	− 0.20	1.05	− 0.68	− 0.28	**1.33**
5	**1.38**	0.51	0.29	0.45	1.24
6	**1.45**	0.97	− 0.16	0.70	0.78
7	0.37	**0.91**	0.46	− 0.72	0.88
8	− 0.06	0.18	− 0.50	0.99	**3.29**
9	0.16	**1.12**	− 0.27	− 0.10	0.08
10	− 0.83	− **0.08**	− 0.97	− 1.81	− 0.49
11	− 0.96	**1.74**	− 1.42	1.03	− 2.27
12	0.24	**0.91**	− 0.37	0.19	0.14
13	− 0.67	**1.23**	0.57	− 1.14	− 0.88
14	− 0.96	**0.63**	− 1.01	− 0.51	− 1.26
15	− 0.03	**1.36**	− 0.03	− 0.10	− 0.75
16	− 0.31	− 1.61	− 0.99	− 1.13	**2.13**
17	− 0.39	**1.02**	0.24	− 1.84	− 0.26
18	− 0.25	− 0.06	**0.34**	− 1.06	− 0.63
19	− 0.61	− 0.63	− **0.07**	− 1.36	− 0.62
20	0.67	0.83	− 0.70	− 0.23	**0.83**
21	0.89	0.76	− 0.32	0.66	**1.12**
22	**1.14**	− 1.40	− 0.90	− 0.16	1.06
23	− 0.79	− 1.67	− 0.71	**2.38**	− 0.22
24	**1.21**	0.72	− 0.64	− 0.59	− 0.22
25	**1.92**	0.81	0.08	0.62	− 0.15
26	− 0.07	− 1.52	**1.07**	0.92	− 1.27
27	− 0.78	**1.24**	− 0.43	0.31	− 1.07
28	− 0.11	**0.58**	0.11	− 0.37	− 0.01
29	0.02	− 1.25	**1.16**	− 0.29	− 1.54
30	− 0.95	− 1.25	0.72	**0.99**	− 0.47
32	− 0.48	− 1.38	0.37	**1.29**	− 0.68
33	− 0.83	− 1.56	− 1.10	**0.28**	− 0.38
34	− 0.94	**0.62**	0.39	− 0.40	0.23
35	**0.73**	0.37	− 0.77	− 0.12	0.46
36	− 0.12	0.65	0.31	0.25	**0.92**
37	**1.51**	− 0.19	0.60	− 0.04	− 1.07
38	**0.64**	− 1.19	− 0.74	0.15	− 1.32
39	− 0.68	0.54	− 0.70	**0.55**	− 0.67
40	**1.48**	0.89	− 0.49	− 0.68	− 0.29
41	− 0.73	− 1.68	− 0.66	− **0.12**	− 0.79
42	− 0.81	0.92	**1.56**	0.93	0.30
43	− 0.36	0.72	− 0.01	**0.77**	0.13
44	− 1.10	− 1.37	**2.37**	1.95	− 0.28
45	− 0.95	**1.55**	0.37	− 0.51	− 0.66
46	0.31	0.55	**2.03**	0.28	0.34
47	**1.91**	0.01	0.89	0.60	− 1.66
48	**3.11**	− 0.08	− 0.75	1.10	− 1.42
49	− 1.00	− **0.27**	− 1.18	− 0.54	− 0.44

Table 10 (continued)

| Sub-watersheds	Factor 1 Magnitude factor	Factor 2 Relief factor	Factor 3 Drainage composition factor	Factor 4 Dissection intensity factor	Factor 5
50	− 0.87	− 0.55	− 2.14	**1.70**	− 0.69
51	0.91	0.72	− 0.56	**1.00**	0.09
52	− 0.86	**0.90**	− 0.76	0.20	− 0.62
53	− 1.07	− 1.18	0.52	**2.03**	0.20
54	− 0.66	− 0.64	− 1.05	− 0.79	**0.13**
55	− 0.68	− 0.03	− 0.69	− 0.39	**0.47**
56	− 0.98	− 0.80	− 0.23	**0.12**	− 0.83
57	− 0.71	**1.52**	− 0.03	− 0.63	− 0.53
58	− 0.41	− 0.53	− 0.38	− 1.65	**1.20**
59	− 1.13	− 0.84	− **0.24**	− 1.40	− 0.47
60	− 0.78	1.01	− 0.81	**1.04**	− 1.66
61	0.63	− 0.41	**1.05**	− 2.30	− 0.03
62	− 0.28	− 0.51	**0.96**	− 1.46	− 0.04
63	**0.79**	0.47	− 0.53	0.23	0.34
64	− 0.51	**1.47**	0.68	0.85	0.58
65	− 0.82	0.60	− 0.36	**1.02**	0.73
66	− 0.24	**0.47**	0.20	− 1.85	0.00
67	**1.14**	0.50	− 0.60	0.00	0.15
68	− 0.50	− 0.45	0.47	− 0.74	**0.51**
69	− 0.56	**0.35**	− 1.14	− 0.90	− 0.23
70	− 0.60	1.58	0.52	**2.16**	− 0.91
71	0.20	− 1.38	− 0.08	0.23	**2.36**
72	− 0.96	− 0.90	− 1.37	− **0.11**	− 0.16
73	0.25	0.99	**1.31**	0.11	0.03
74	**0.78**	0.51	− 0.72	− 0.81	0.31
75	− 0.51	0.56	0.43	0.45	**0.89**
76	− 0.91	0.32	**1.48**	− 1.81	0.43
77	**1.30**	0.28	− 0.03	− 0.14	0.77
78	− 1.05	0.94	− 0.17	**1.61**	0.73
79	− 1.20	0.28	**0.99**	0.78	0.88
80	− 0.94	**0.60**	− 1.47	− 0.46	− 0.11
81	− 0.67	− 1.30	**1.60**	1.37	1.43
82	− 0.26	− 0.33	**4.10**	0.08	0.05
83	**1.77**	0.18	− 0.64	0.24	1.00
84	0.91	− 0.47	**2.53**	− 0.43	− 0.91
85	− 1.14	− 1.36	− 1.34	**0.02**	− 0.43
86	0.68	− 0.88	**1.43**	− 0.89	− 1.29
87	− 0.05	− 1.66	− 0.86	− **0.01**	− 0.47
88	0.06	− 1.72	**0.70**	− 1.97	− 0.24
89	**0.25**	− 1.91	− 1.25	− 1.42	0.08
90	**2.06**	− 1.12	− 0.23	0.22	− 1.42
91	**1.83**	− 0.88	0.90	0.50	− 1.45
93	− 1.00	− 1.83	0.10	0.87	**2.32**
94	**3.39**	− 1.18	− 1.13	0.82	0.02
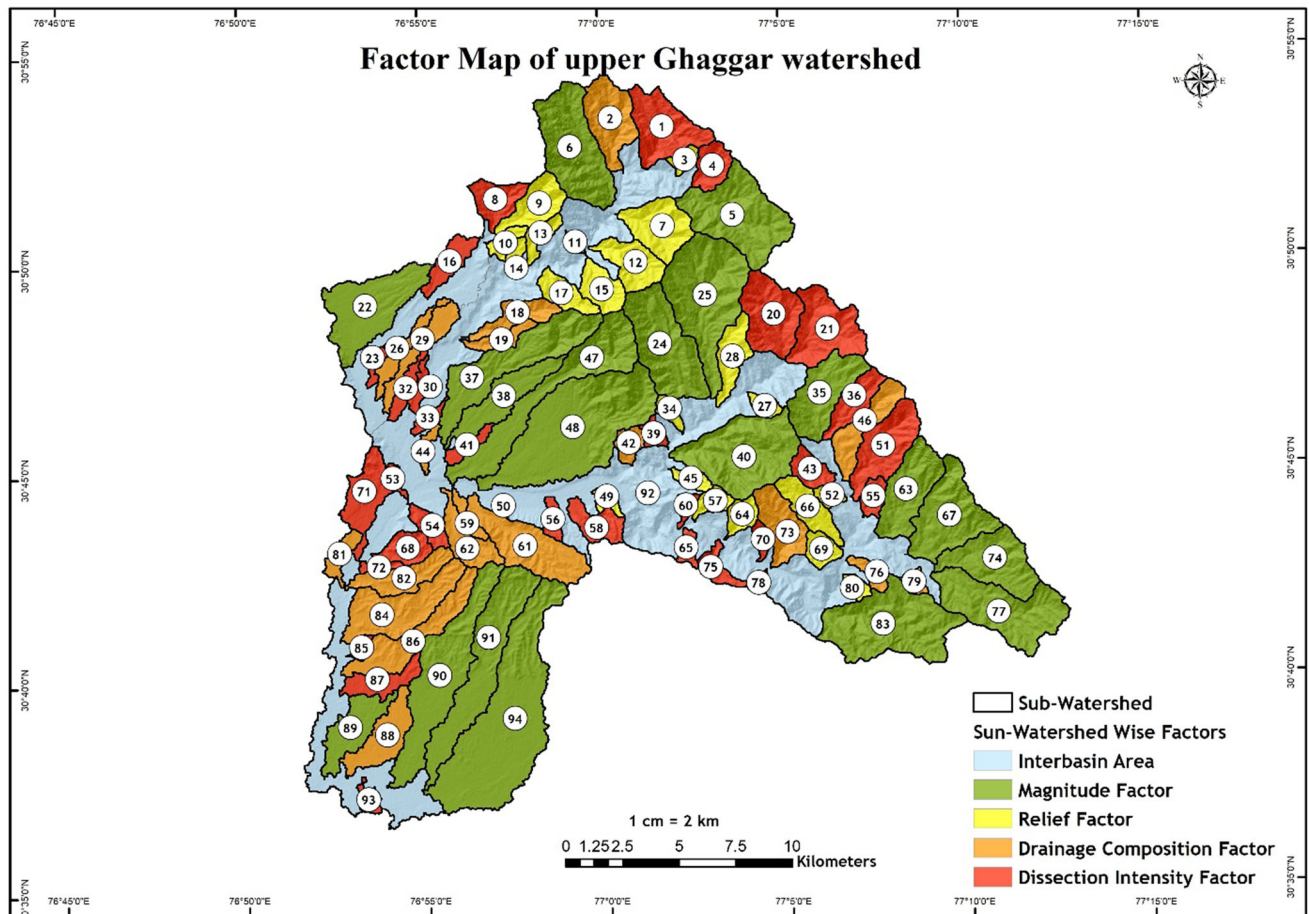
Fig. 5 Dominating factor scores prevalent within the upper Ghaggar watershed

transport. Therefore, the sub-watersheds having higher R_{hp} having peak discharges resulting in higher erosion. The mean values of D_{is}, R_n, and S of the sub-watersheds falling in different clusters range from 0.16 to 0.25, 0.30 to 0.88, and 7.92 to 22.67, respectively, as shown in Table 11. Higher D_{is}, R_n, and S values correspond to sub-watersheds falling under Cluster 1, while low values of D_{is}, R_n, and S correspond to sub-watersheds falling under Cluster 3. The higher values of D_{is}, R_n, and S indicate steeper slopes, higher dissection, less overland flow due to lower time of

concentration resulting in higher possibilities of flash floods and more erosion than sub-watersheds having lower D_{is}, R_n, and S.

The average values of all the Cluster's priorities 1, 2 and 3 are 1.77, 2.07, and 2.16, respectively. The spatial distribution of the priority-wise sub-watersheds is shown in Fig. 7. The Cluster 1 sub-watersheds are with the highest priority, followed by Cluster 2, and finally, the least priority sub-watersheds falling under Cluster 3.
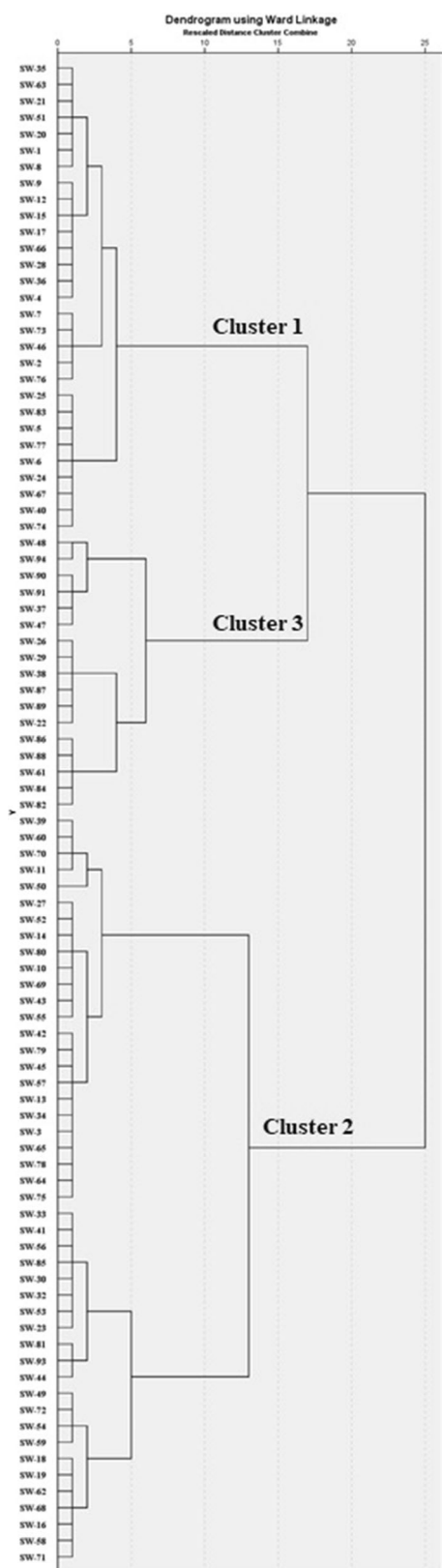
Fig. 6 Dendrogram of the clustered sub-watersheds in upper Ghaggar watershed

Conclusions

The upper Ghaggar watershed, forming the part of the lower Shivaliks, is the most fragile ecosystem due to its highly erodible soil structures. Therefore, prioritization at the sub-watershed scale can help the decision/policy makers to address the soil erosion problems eminent within a specific region of the upper Ghaggar watershed. The area is also prone to flooding during the monsoon season, hence water conservation structures such as check dams, gabion walls, etc., can be constructed in the downstream of the basin which will reduce the potential of flash floods and will increase the groundwater recharge and utilization of surface water for irrigation. In the present study, watershed prioritization of the erosion susceptible areas has been achieved by synergistically fusing the remote sensing and geographical information system techniques backed with statistical techniques such as correlation analysis, principal component analysis, and hierarchical clustering analysis. Twenty-seven linear, areal, and relief morphometric parameters were chosen for ninety-two sub-watersheds, where thirty-two are fourth-order, and sixty are third-order. The Principal Component Analysis applied on the inter-correlation matric resulted in five significant components representing the entire dataset. These five principal components were grouped into four factors, i.e., magnitude factor (PC1), relief factor (PC2), drainage composition factor (PC3), and Dissection intensity factor (PC4 & PC5). The Upper Ghaggar watershed is primarily controlled by magnitude and relief factors since these together explain 67.66% of the total variance. Because of the larger number of sub-watersheds involved in the present study, hierarchical clustering techniques based on Ward's algorithm resulted into three prime clusters, i.e., high, medium, and least priority showing considerable spatial variability. Out of ninety-two, 29, 46, and 17 sub-watersheds fall in high, medium, and least priorities clusters. The results of this study provide vital information to the planners and decision-makers regarding drainage morphometry utilizing the geospatial techniques, which can act as a base criterion for implementing soil erosion conservation strategies (such as planting vegetation, hillshade terracing) and assessment of surface/groundwater recharge potential and flash flood mitigation (by constructing water harvesting structures such as check dams, gabion structures, etc.).

Table 11 Results of morphometric parameters derived based on hierarchical cluster analysis method for the 92 sub-watersheds

Linear characteristics

Cluster ID	N_u	SL_u	MSL_u	$MSLR_u$	R_b	ρ
Cluster 1	57.90 (2)	20.74 (2)	4.51 (2)	3.10 (2)	4.25 (2)	0.73 (1)
Cluster 2	13.17 (3)	4.97 (3)	1.92 (3)	2.52 (3)	3.03 (3)	0.83 (2)
Cluster 3	72.76 (1)	36.35 (1)	8.00 (1)	4.35 (1)	4.98 (1)	0.86 (3)

Shape characteristics

Cluster ID	L_k	F_f	R_e	I_e	R_c
Cluster 1	1.73 (2)	0.46 (2)	0.76 (2)	2.62 (1)	0.31 (1)
Cluster 2	1.40 (3)	0.56 (1)	0.85 (1)	2.75 (2)	0.29 (2)
Cluster 3	1.80 (1)	0.44 (3)	0.75 (3)	6.35 (3)	0.17 (3)

Drainage characteristics

Cluster ID	D_d	S_f	D_t	D_i	I_f
Cluster 1	2.66 (3)	11.43 (3)	0.72 (3)	4.58 (1)	32.95 (3)
Cluster 2	3.16 (2)	11.97 (1)	0.60 (1.5)	3.92 (2)	41.01 (2)
Cluster 3	3.58 (1)	11.44 (2)	0.51 (1.5)	3.32 (3)	43.20 (1)

Relief characteristics

Cluster ID	R_a	R_r	R_{hp}	D_{iS}	R_n	S
Cluster 1	1321.11 (1)	333.95 (1)	2.25 (2)	0.25 (1)	0.88 (1)	22.67 (1)
Cluster 2	792.68 (2)	215.28 (2)	3.22 (1)	0.24 (2)	0.63 (2)	15.36 (2)
Cluster 3	512.57 (3)	94.39 (3)	0.37 (3)	0.16 (3)	0.30 (3)	7.92 (3)
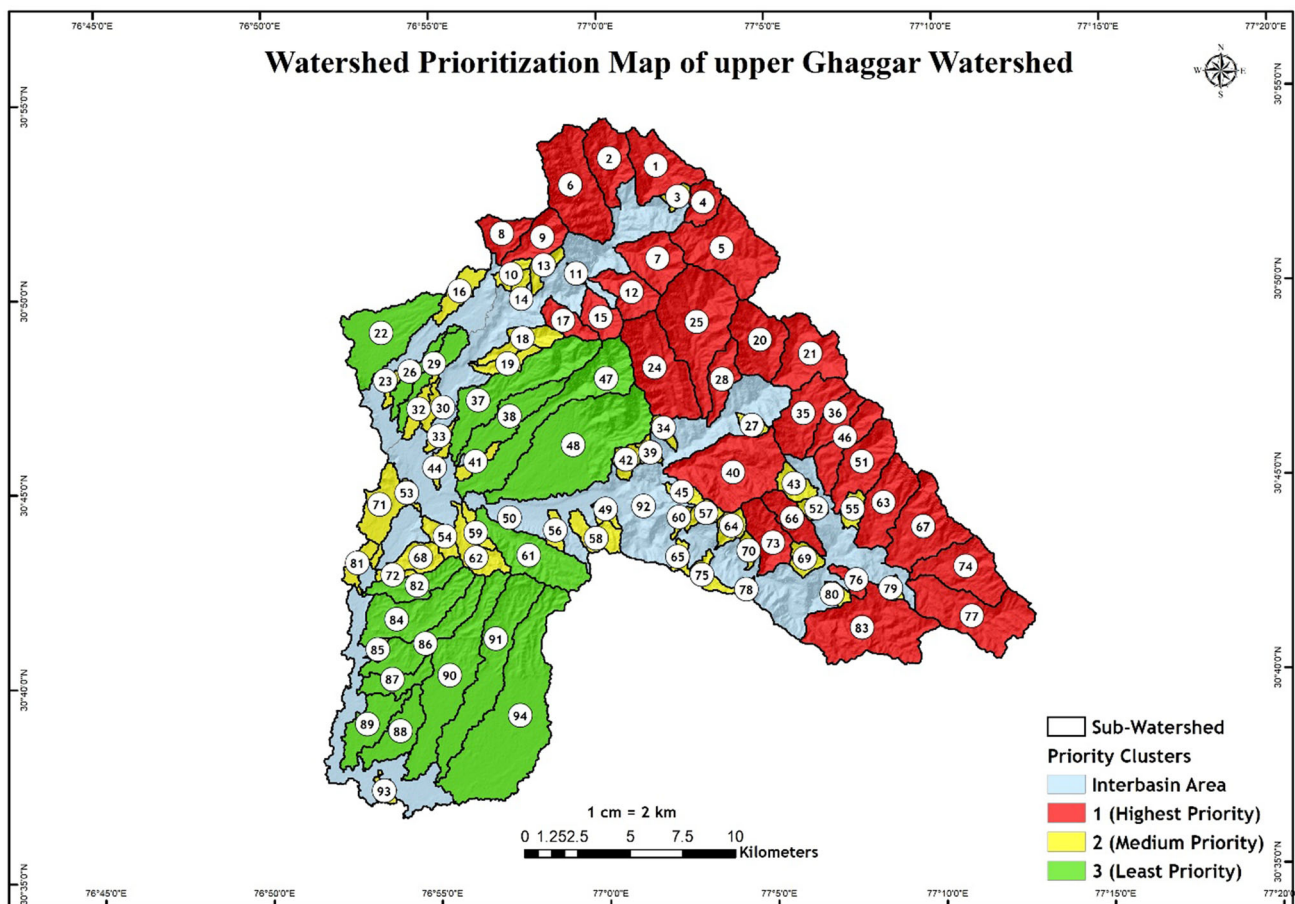
Fig. 7 Final prioritized sub-watersheds of upper Ghaggar watershed

Declarations

References

Abdeta, G. C., Tesemma, A., Tura, A. L., & Atlabachew, G. H. (2020). Morphometric analysis for prioritizing sub-watersheds and management planning and practices in Gidabo Basin, Southern Rift Valley of Ethiopia. *Applied Water Science*. https://doi.org/10.1007/s13201-020-01239-7

Arefin, R., Mohir, M., & Alam, J. (2020). Watershed prioritization for soil and water conservation aspect using GIS and remote sensing: PCA-based approach at northern elevated tract Bangladesh. *Applied Water Science*. https://doi.org/10.1007/s13201-020-1176-5

Bhat, M. S., Alam, A., Ahmad, S., Farooq, H., & Ahmad, B. (2019). Flood hazard assessment of upper Jhelum basin using morphometric parameters. *Environmental Earth Sciences*. https://doi.org/10.1007/s12665-019-8046-1

Bothale, R. V., Bothale, V. M., & Sharma, J. R. (1997). Delineation of eco watersheds by integration of remote sensing and GIS techniques for management of water and land resources. In D. Fritsch, M. Englich, & M. Sester (Eds.), *ISPRS commission IV symposium on GIS—between visions and applications. 32/4*. IAPRS.

Chandniha, S. K., & Kansal, M. L. (2017). Prioritization of sub-watersheds based on morphometric analysis using geospatial technique in Piperiya watershed, India. *Applied Water Science, 7*, 329–338. https://doi.org/10.1007/s13201-014-0248-9

Chauhan, N., Kumar, V., & Paliwal, R. (2020a). Land capability assessment of Ghaggar river basin using integrated remote sensing and geographical information system approach—A Case Study. *Annals of Plant and Soil Research, 22*(4), 367–372. https://doi.org/10.47815/apsr.2020a.10006

Chauhan, N., Kumar, V., & Paliwal, R. (2020b). Quantifying the impacts of decadal landuse change on the water balance components using soil and water assessment tool in Ghaggar river basin. *SN Applied Sciences*. https://doi.org/10.1007/s42452-020-03606-0

Chen, C. W., & Zebker, H. A. (2000). Network approaches to two-dimensional phase unwrapping: Intractability and two new algorithms. *Journal of the Optical Society of America, 17*(3), 401–414.

Chorley, R. J., Malm, D. E., & Pogorzelski, H. A. (1957). A new standard for estimating drainage basin shape. *American Journal of Science, 255*, 138–141.

Chow, V. T. (1964). *Handbook of applied hydrology*. McGraw-Hill Book Co., Inc.

Dadhwal, V. K., Sharma, J. R., Bera, A. K., Paliwal, R., Chauhan, N., & Shirsath, P. B., et al. (2012). *River Basin Atlas of India*. Central Water Commision and Regional Remote Sensing Centre-West, NRSC, ISRO.

F.A.O. (1985). *Watershed development with special reference to soil and water conservation*. Rome: F.A.O Soil Bulletin 44.

Falkenmark, M., & Lindh, G. (1976). *Water for a starving world*. Westview Press.

Faniran, A. (1968). The index of drainage intensity—A provisional new drainage factor. *Australian Journal of Scienec, 31*, 328–330.

Gajbhiye, S., & Mishra, S. K. (2012). Application of NRSC-SCS curve number model in runoff estimation using RS & GIS. In *IEEE-international conference on advances in engineering, science and management (ICAESM-2012)* (pp. 346–352). IEEE.

Gardiner, V. (1975). *Drainage basin morphometry*. British Geomorphological Research Group.

Goldstein, R. M., Zebker, H. A., & Werner, C. L. (1988). Satellite radar interferometry: Two-dimensional phase unwrapping. *Radio Science, 23*(4), 713–720.

Gopinath, G., Nair, A. G., Ambili, G. K., & Swetha, T. V. (2016). Watershed prioritization based on morphometric analysis coupled with multi criteria decision making. *Arabian Journal of Geosciences, 9*, 1–17. https://doi.org/10.1007/s12517-015-2238-0

Horton, R. E. (1932). Drainage-basin Characteristics. *Transactions, American Geophysical Union, 13*(1), 350–361.

Horton, R. E. (1945). Erosional development of streams and their drainage basins; Hydrophysical approach to quantitative morphology. *Geological Society of America Bulletin, 56*(3), 275–370.

Jhariya, D. C., Kumar, T., & Pandey, H. K. (2020). Watershed prioritization based on soil and water hazard model using remote sensing, geographical information system and multi-criteria decision analysis approach. *Geocarto International, 35*, 188–208.

Mangan, P., Haq, M., & Baral, P. (2019). Morphometric analysis of watershed using remote sensing and GIS—A case study of Nanganji River Basin in Tamil Nadu, India. *Arabian Journal of Geosciences, 12*(6), 14.

Mather, P. M., & Doornkamp, J. C. (1970). Multivariate analysis in geography with particular reference to drainage-basin morphometry. *Transactions of the Institute of British Geographers, 51*, 163–187. https://doi.org/10.2307/621768

Melton, M. A. (1958). Correlation structure of morphometric properties of drainage systems and their controlling agents. *The Journal of Geology, 66*(4), 442–460.

Meshram, S. G., & Sharma, S. K. (2017). Prioritization of watershed through morphometric parameters: A PCA-based approach. *Applied Water Scinece*. https://doi.org/10.1007/s13201-015-0332-9

Miller, V. C. (1953). *A Quantitative geomorphic study of drainage basin characteristics in the Clinch Mountain area, Virginia and Tennessee*. Columbia University, Department of Geology.

Mishra, S. K., Gajbhiye, S., & Pandey, A. (2013). Estimation of design runoff curve numbers for Narmada watersheds (India). *Journal of Applied Water Engineering and Research, 1*(1), 69–79. https://doi.org/10.1080/23249676.2013.831583

Narayana, R. D., & Babu, R. (1983). Estimation of soil erosion in India. *Journal of Irrigation and Drainage Engineering, 109*(4), 419–434.

Nautiyal, M. D. (1994). Morphometric analysis of a drainage basin using aerial photographs: A case study of Khairkuli basin, district Dehradun, U.P. *Journal of the Indian Society of Remote Sensing, 22*, 251–261. https://doi.org/10.1007/BF03026526

Nir, D. (1957). The ratio of relative and absolute altitudes of Mt. Carmel: A contribution to the problem of relief analysis and relief classification. *Geographical Review, 47*(4), 564–569.

Nookaratnam, K., Srivastava, Y. K., Rao, V. V., Amminedu, E., & Murthy, K. R. (2005). Check dam positioning by prioritization of micro-watersheds using SYI model and morphometric analysis—Remote Sensing and GIS perspective. *Journal of the Indian Society of Remote Sensing, 33*(1), 2–38.

Obeidat, M., Awawdeh, M., & Al-Hantouli, F. (2021). Morphometric analysis and prioritisation of watersheds for flood risk management in Wadi Easal Basin (WEB), Jordan, using geospatial technologies. *Journal of Flood Risk Management.* https://doi.org/10.1111/jfr3.12711

Oldeman, L. R. (1991). *Global extent of soil degradation.* ISRIC: Wageningen.

Patton, P. C., & Baker, V. R. (1976). Morphometry and floods in small drainage basin subject to diverse hydrogeomorphic controls. *Water Resource Research, 12*(5), 941–952.

Prasad, R. K., Mondal, N., Banerjee, P., Nandakumar, M. V., & Singh, V. S. (2008). Deciphering potential groundwater zone in hard rock through the application of GIS. *Environmental Geology, 55,* 467–475. https://doi.org/10.1007/s00254-007-0992-3

Prieto-Amparán, J. A., Pinedo-Alvarez, A., Vázquez-Quintero, G., Valles-Aragón, M. C., Rascón-Ramos, A. E., Martinez-Salvador, M., & Villarreal-Guerrero, F. (2019). A multivariate geomorphometric approach to prioritize erosion-prone watersheds. *Sustainability.* https://doi.org/10.3390/su11185140

Rai, P. K., Mohan, K., Mishra, S., Ahmad, A., & Mishra, V. N. (2017). A GIS-based approach in drainage morphometric analysis of Kanhar River Basin, India. *Applied Water Science, 7,* 217–232. https://doi.org/10.1007/s13201-014-0238-y

Rai, P. K., Chandel, R. S., Mishra, V. N., & Singh, P. (2018). Hydrological inferences through morphometric analysis of lower Kosi river basin of India for water resource management based on remote sensing data. *Applied Water Sciences, 8*(15), 1–16. https://doi.org/10.1007/s13201-018-0660-7

Rao, D. P. (2000). Role of remote sensing and geographical information system in sustainable development. In *International archives of photogrammetry and remote sensing* (Vol. XXXIII, pp. 1231–1251). Amsterdam: ISPRS.

Schumm, S. A. (1956). Evolution of drainage systems and slopes in badlands at perth amboy, New Jersey. *Geological Society of America Bulletin, 67*(5), 597–646.

Siddiqui, R., Said, S., & Shakeel, M. (2020). Nagmati River sub-watershed prioritization using PCA, integrated PCWS, and AHP: A case study. *Natural Resources Research, 29,* 2411–2430. https://doi.org/10.1007/s11053-020-09622-6

Singh, S. (1976). On the quantitative parameters for the computation of drainage density, texture and frequency: A case study of a part of the Ranchi Plateau. *National Geographer, 10*(1), 21–31.

Singh, S., & Singh, M. C. (1997). Morphometric analysis of Kanhar river. *National Geographical Journal of Lndia, 43*(1), 31–43.

Singh, G., Babu, R., Narain, P., Bhushan, L. S., & Abrol, I. P. (1992). Soil erosion rates in India. *Journal of Soil and Water Conservation, 47*(1), 97–99.

Smith, G. H. (1935). The relative relief of Ohio. *Geographical Review, 25,* 247–248.

Space Application Centre. (2016). *Desertification and land degradation atlas of India (Based on IRS AWiFS data of 2011–13 and 2003–05).* Space Applications Centre, ISRO.

Stoddart, D. R. (1965). The shape of atolls. *Marine Geology, 3*(5), 369–383.

Strahler, A. N. (1952). Hypsometric (area-altitude curve) analysis of erosional topography. *Geological Society of America Bulletin, 63*(11), 1117–1142.

Strahler, A. N. (1953). Revisions of Horton's quantitative factors in erosional terrain. *Transactions American Geophysical Union, 34,* 345.

Strahler, A. N. (1957). Quantitative analysis of watershed geomorphology. *Transactions, American Geophysical Union, 38*(6), 913–920.

Strahler, A. N. (1964). Quantitative geomorphology of drainage basins. In V. T. Chow (Ed.), *Handbook of applied hyrology* (pp. 4–11). McGraw Hill Book Company.

Thakkar, A. K., & Dhiman, S. D. (2007). Morphometric analysis and prioritization of miniwatersheds in Mohr watershed, Gujarat using remote sensing and GIS techniques. *Journal of Indian Society of Remote Sensing, 35,* 313–321. https://doi.org/10.1007/BF02990787

Waiyasusri, K., & Chotpantarat, S. (2020). Watershed prioritization of Kaeng Lawa sub-watershed, Khon Kaen Province using the morphometric and land-use analysis: A case study of heavy flooding caused by tropical storm podul. *Water.* https://doi.org/10.3390/w12061570

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of American Statistical Association, 58*(301), 236–244. https://doi.org/10.1080/01621459