



Determining the spatio-temporal response of downstream coarse sediment sorting process in the Chel river (North Bengal, India) using cluster analysis

Debarshi Ghosh¹ · Snehasish Saha²

Received: 23 July 2020 / Accepted: 17 October 2020 / Published online: 3 November 2020
© Springer Nature Switzerland AG 2020

Abstract

The fluvial transportation efficiency related to rainstorm events is important to understand the variability of sediment sorting process in the channel. The supply limitation of different size of sediment grades in downstream reveals the nature of geomorphological scale of response (spatio-temporal). Pebble count method has been attempted here for determination of d_{50} (median) particle size associated with frequency distribution of sediment samples from 13 study reaches on the basis of standard sediment grade scale of Wentworth and modified Udden-Wentworth grain scale seasonally. Intra temporal variability (IATV) and inter temporal variability (IETV) results were categorized under hierarchical cluster analysis to view the sediment movement downstream from minimum distance algorithm. On the piedmont, the bed configuration is mostly boulder infested of heavier size class (> 256 mm) and that remains greater than 86% in pre-monsoon condition. It comes out on the basis of random mixing of sediment load (> 2 mm) while transporting as discharge. In Chel basin, the process of sediment load dispersal indicates less heterogeneity in the sediment sorting process towards downstream. The haphazard distribution of very coarse to medium boulders (4096–512 mm) up to reach 6 indicates the limit of channel competency. The paper seeks to know the seasonal variability in the sediment dispersal process i.e. finding the reach wise variability of the sediment mixing process downstream and the sediment supply limits of median particle size.

Keywords Sediment sorting · Frequency distribution curve · Normal probability distribution · Hierarchical cluster · ANOVA

Introduction

The main goal of cluster analysis is to build homogeneous groups from a geographical area according to certain variables (De Carvalho and Lechevallier 2009; Dutta et al. 2019, Dutta and Das 2019). Statistical parameters like the mean, median, standard deviation, kurtosis, and skewness have been widely used to characterize grain size distribution

within the channel bed and clustering finds the similar distributions. Most of the studies have employed the method to group a limited number of grain size distribution parameters like the mean and standard deviation, and few studies have taken advantage of the entire distribution (Zhou et al. 2015; Nelson et al. 2014; Ordóñez et al. 2016). Thus, both the process helps to infer variations in hydrodynamic conditions, eolian activity, and sediment sources (Fournier et al. 2014; Zhang et al. 2018). The frequency distribution and mixing process of various grain sizes as bed load can be expressed through percentile-median diameter (CM) pattern (Passega 1964), standard deviation method (Boulay et al. 2003), curve fitting models (Wu et al. 2020; Paterson and Heslop 2015) and bayesian method (Yu et al. 2016) of sediment mixing process. The cluster analysis groups similar conditions of grain size mixing process with end-member modeling (Varga et al. 2019). The cluster analysis is also an important spatial analysis tool. The modern clustering processes using space-time data comes from different shapes

✉ Debarshi Ghosh
wetlanddeb@gmail.com

Snehasish Saha
snehasishsahanbu1979@gmail.com

¹ Department of Geography, Dhupguri Girls' College (Affiliated to University of North Bengal), Dhupguri, West Bengal, India

² Department of Geography and Applied Geography, University Of North Bengal, Darjeeling, West Bengal, India

and utilities viz. partitioning methods (K-means and K-centroid technique algorithms), hierarchical methods (Chameleon, Ward's Method, Nearest Neighbor algorithms), density based approaches (OPTICS and SNN algorithms) and grid based methods (STING and CLIQUE algorithms) (Han et al. 2009; Lamb et al. 2016, 2020). The conformation of a prior hypothesis can be done by traditional hierarchical clustering. In contrast, exploratory methods rely on unsupervised machine learning algorithms to identify groups of data in n -dimensional space by using a similarity or dissimilarity metric (Everitt et al. 2011). The exploratory methods of clustering determine the pattern of unknown spatio-temporal variability within the nature. There exists the problem of over-estimation in clustering for linearly constrained point events in confirmatory case (Okabe and Yamada 2001; Lamb et al. 2016). The problem occurs when the euclidean distance underestimate the distance between points. Thus, a network distance metric is preferred. In modern geomorphological studies the clustering approach is being popularized to understand the process of sediment dispersal, channel migration and channel responses in bar formation (Tal et al. 2004; Church 2010; Hayashi et al. 2012; Constantine et al. 2014; Corenblit et al. 2015). Cluster analysis applied to sedimentary grain-size data is usually used for stratigraphic subdivision; and determination of the sources of sediments and sediment mixing process in the river beds (Donato et al. 2009; Liu et al. 2017). The cluster analysis also can be used to show the nature of sediment mixing and dispersal process in the high fluvial energy environments. The process of sediment load dispersal in the high fluvial energy transportation regimes brings frequent changes in the channel planform (Rinaldo 1991; Yang and Shi 2019; Walsh and Nittrouer 2009). The rivers of piedmont in northern part of West Bengal do not keep pace with rapid sediment gradation in downstream due to variable trend of sediment delivery from the upper catchment (Starkel et al 2008). The rain storm events (1992, 1996, 1998 and 2000) of Darjeeling sub-Himalaya created mass-flow of large debris which gradually spread over the fan surface and left ferruginous coats while rootlet debris concentrated on the surface piedmont (Starkel et al. 2008; Mandal and Sarkar 2016; Karen et al. 2015; Blair and Mcpherson 1999; Doeglas 1968). The transport efficiency of the channel related to discharge and bed roughness causes variability of sediment sorting process downstream (Kondlof and Li 1992; Kleinhans 2002; Luo et al. 2013; Ma et al. 2017). The sediment sorting from outlet to sink is largely dependent on the channel connectivity with distributaries in the upper catchment which control the flux of sediment delivery to the channel (Syvitski et al. 2005). In Chel basin, the changing monsoon rainstorm intensity, channel runoff and morphological controls make the channel

limited in terms of both supply and capacity (Hickin 1995). These limitations bring high variability in sediment movement and gradual sorting process. This paper proposes a spatio-temporal clustering approach which pivoted on the agglomerative hierarchical clustering to identify similar reaches in terms of median particle size and the downstream limit of its distribution. The approach of space-time hierarchical clustering incorporates location, time, and attribute information to identify the groups across a nested structure reflective of a hierarchical interpretation of scale (Manson 2008). The process of mixed sediment sorting and their similarity in downstream has been understood from the geomorphological (spatio-temporal) scale responses (Boer 1992). The approach reveals the role of continuously interacting complex set of fluvial process and its competency in understanding sediment sorting downstream in pre- and post-monsoon seasons.

The study area

The river Chel originates from the Pankhashari reserved forest of Kalimpong district at an elevation of 2484 above mean sea level (Fig. 1). The river flows through the piedmont tract and flat rolling alluvial surface combining another two streamlets namely Fagu and Manzing khola from the left. The mountain catchment covers 97 km² area and ends at the mountain outlet near Patharjhora tea garden (Jalpaiguri district). From outlet, the course of 27 km length finally debouches into the Tista river at Rajadanga gram panchayet (Mal block). The total basin area is 317 km². The heterogeneous sediment load on the piedmont surface is the composition of lingtse granite, gneissic boulders, undifferentiated Siwalik formation, Pleistocene rocks, and fine Quaternary sand deposition (Fig. 1). It originates from active thrust zone (Main Frontal Thrust) in upper catchment and concentrates in the channel from rain wash and surface runoff (Sarkar 2012; Chakraborty and Nag 2015).

Methodology

The fluvial mechanics bound to act within the balance of erosion and deposition. The displacement of material by fluvial processes (drilling, abrasion, corrosion, helical flow action) acts as erosion. The amount and size of sediments which move in sliding, suspension, rolling and saltation depend on the three fundamental characteristics of flow viz. channel capacity, channel competency and sediment supply. The methodological framework of the study has been presented in this section (Fig. 2).

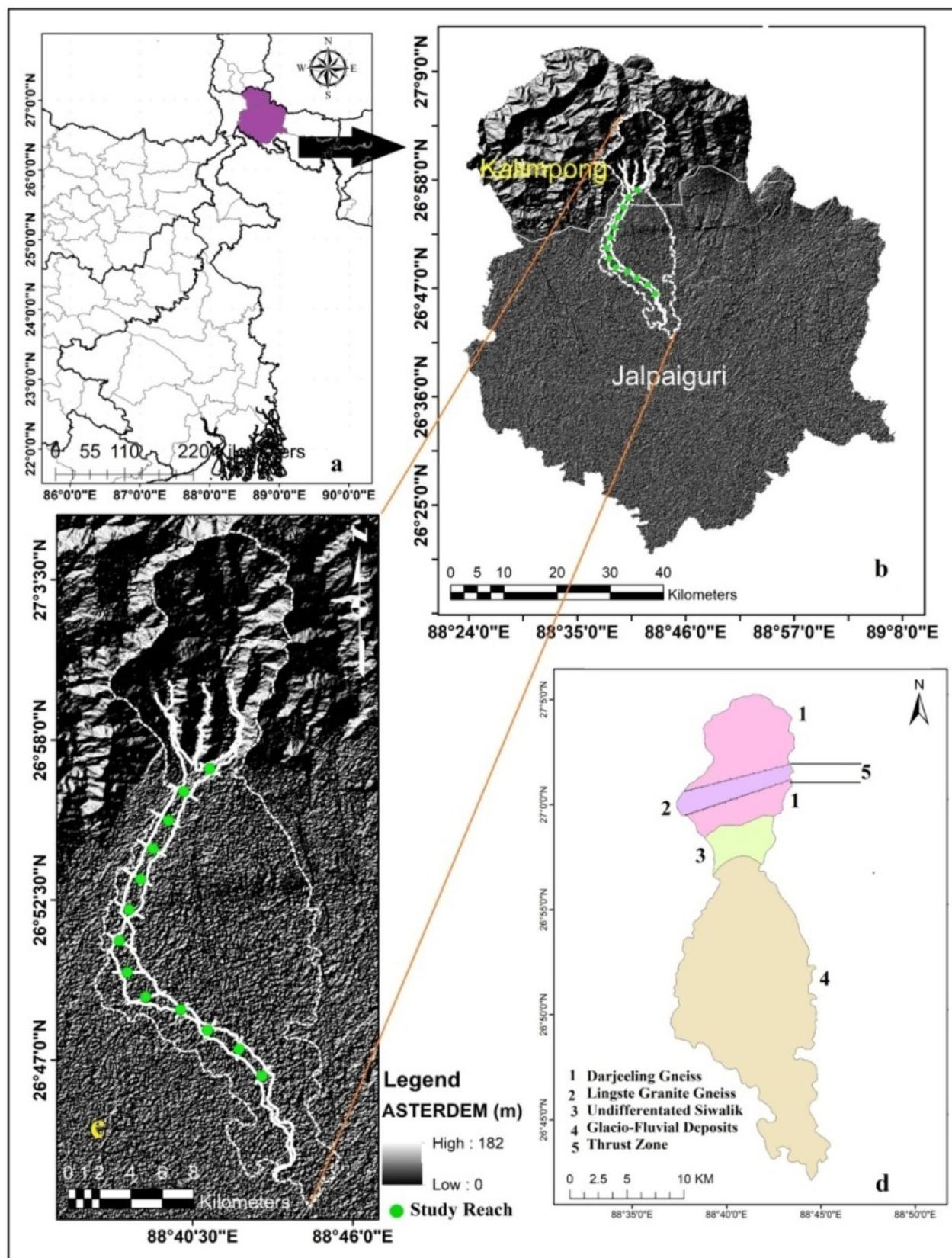


Fig. 1 Location of the Chel basin

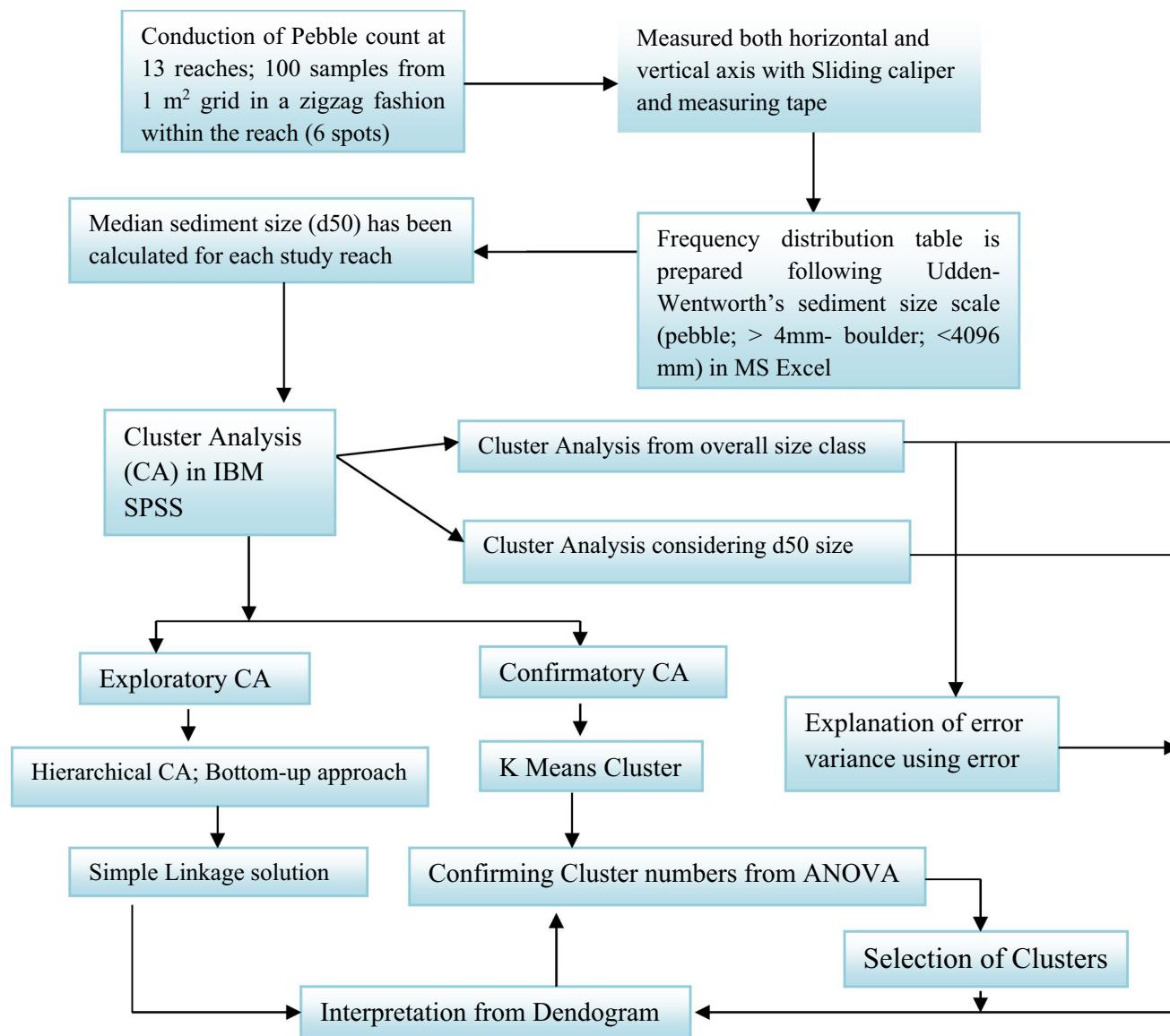


Fig. 2 Flow chart of methodology

Pebble count method

In this technique, 100 random samples (> 2 mm) have been collected from each study reach by placing 1m^2 grid (Plate 1f) frame and placement points are derived by dividing the reach distance in 4 parts (Tamang 2013). The samples have been counted both on near bank and mid channel parts. The study of grid sampling (1m^2) (Wolman 1954; Folk 1954; Julian 1998; López 2017; Bunte and Abt 2001) has been conducted from Patharjhora (mountain outlet) at a regular interval of 2 km up to Rajadanga site (upstream to downstream) to know the nature of sediment sorting and downstream variation according to the fluvial process. The intermediate axis

of each sample has been measured with a sliding caliper in the field (Plate 1g). The Udden–Wentworth sediment grade scale has been considered for generating frequency of respective size classes following pebble count method (Blair and Mcpherson 1999; Wolman 1954; Folk 1954; Julian 1998; Blott and Pye 2012; Terry and Goff 2014). The authors considered the pebble size categories as pebble (2–64 mm), cobble (64–256 mm) and boulder (> 256 mm) following Udden–Wentworth sediment grade scale. The pebble count results are theoretically equivalent to size distributions obtained from bulk samples (Kellerhals and Bray 1971). For the wide alluvial channels, the repetitive sampling is needed to get close match with the theoretical frequency distribution.

Determination of d_{50} particle size

The sediment grain size distribution is a technique to find out the heterogeneous character of sediment dispersal (Zhang et al. 2014a, b, 2018; Qiu et al. 2016). The river maintains sediment sorting process with flow variations in downstream. The term d_{50} represents the median particle size. It is an important indicator in characterization of particle size. Here, the method of d_{50} size determination is taken up to know the dominant d_{50} size in each study reach and its exact fitting in the Udden-Wentworth grade scale category. The d_{50} presents the central sediment size among large scale of size variation. The median particle size variation is considered to know the sediment supply limit and mixing process within the channel.

Frequency distribution of sediment samples and statistical representation

The frequency distribution of sediment samples from 13 study reaches apart from 2 km interval has been studied on the basis of standard sediment grade scale of Wentworth (1922) and modified Udden-Wentworth grain scale both in pre- and post-monsoon conditions. The shape of the Frequency Distribution Curves (FDC) is studied based on the concept of kurtosis to know the tailedness of the distribution within a particular range of sediment grade (Clark 1976; Brasington et al. 2003; Poppe et al. 2004; Gayer et al. 2008). It also indicates the modality or multi-modality of the FDC which obviously relates to sequential role of flood discharge with varying intensity (Jerosch 2013). The left or right tailedness of the FDC is understood on skewness function. The left tailed curve (negative skewness) relates the phases of extreme events in the channel with narrow band of peakedness on the higher values of abscissa. The amount of variance in sampling procedure has been understood from the shape of the distribution keeping resemblance with the theoretical Gaussian distribution (Leys et al. 2005). The heterogeneity of the sample frequency distribution has been visualized from the surface matrix plots (SMP) on a cross transect array from outlet to sink. The pre- and post-monsoon (2015–2017) variability of sediment sample sizes (> 2 mm) has been plotted in bubble diagram.

Algorithmic expression of hierachic cluster analysis (HCA)

Hierarchical cluster (HC) is an algorithm that clubs similar objects into groups, known as clusters. HC decomposes data into a hierarchy of groups; each larger group contains a set of subgroups (Zhang et al. 2016; Lamb et al. 2020).

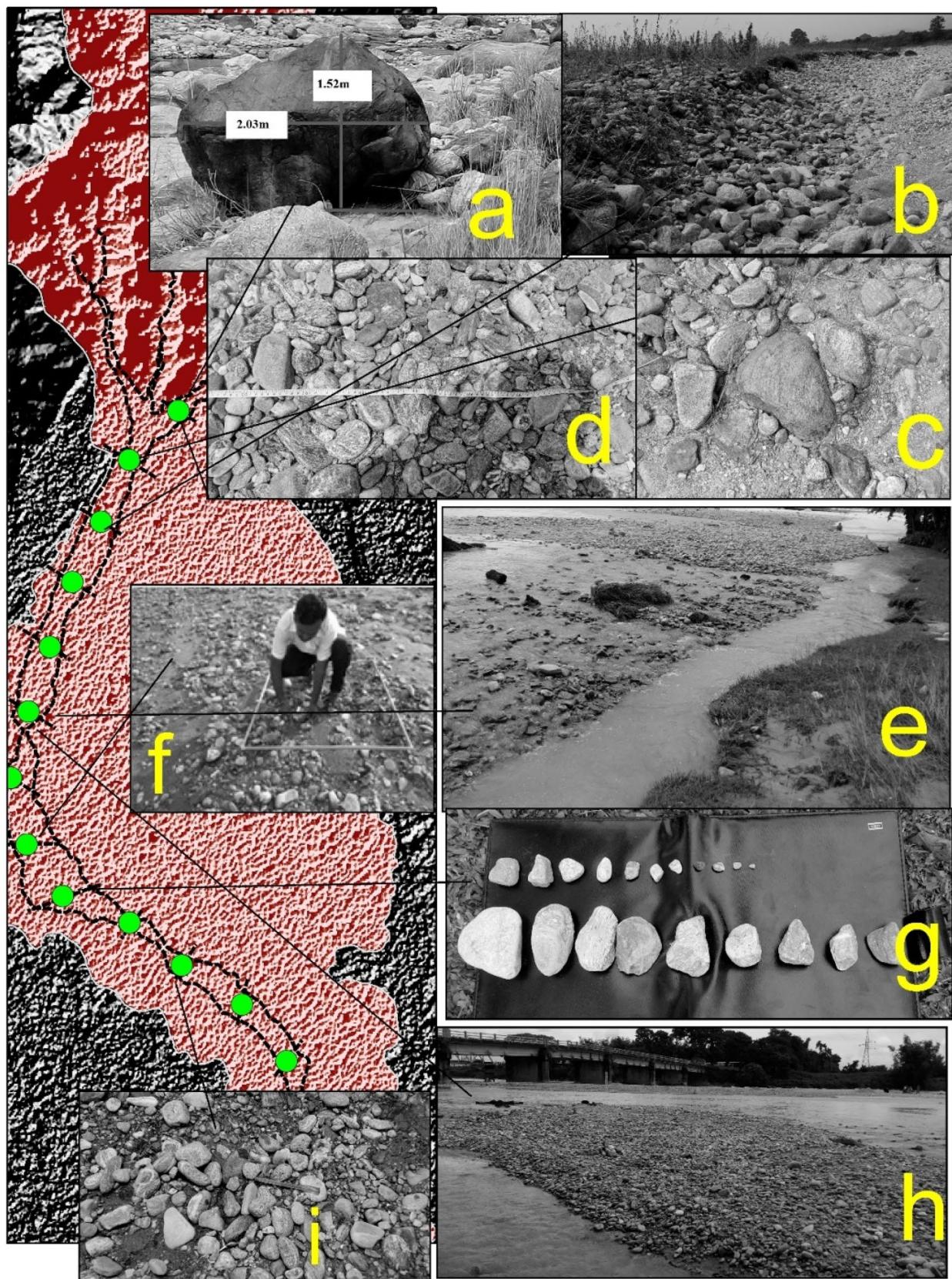
If the approach is started from the bottom with individual observations considered each as a cluster then it becomes aggregated clustering. In the agglomerative approach, there are at least four methods of merging clusters: single-linkage, complete linkage, average group linkage and centroid linkage. The objects within each cluster are broadly similar to each other. The HC analysis involves in solving firstly, the similarity measure that can be used as a scalar distance between different clusters and secondly finding the linkage method that orders the clusters to produce a unique and meaningful solution (Langfelder et al. 2008; Lu et al. 2017; Tonnen et al. 2014; Zhang et al. 2017). The heterogeneity of surface sediment distribution in Chel basin comprises with single linkage solution (SLS) (Dawyndt et al. 2006; Zhang et al. 2019). It connects the two most similar bits of a cluster. The single linkage solution comes under agglomerative hierarchical clustering (AHC) which works on the principle of sequential merging of similar clusters (bottom-up approach). The single linkage methods use the nearest distance between members of clusters. The agglomerative clustering has been selected in this paper because of two reasons: firstly, in agglomerative schedule, it is not necessary to pre-select the number of clusters and secondly, it helps to view clusters at different levels with a hierachic scale. The efficiency of the linkage criteria has been estimated by SLINK or single linkage function. The method is also known as nearest neighbor clustering. The distance between the clusters equals the distance between the elements in each cluster that are remaining nearest from each other. The spatial distance matrix follows the euclidean distance as the movement of d_{50} particle size has been conceptualized on 2-dimensional planner space capturing the distance using x and y coordinate expresses in the equation number 1:

$$D_E(x_i, x_j)^2 = (xi - xj)^2 + (yi - yj)^2. \quad (1)$$

The main steps while performing HC analysis in this paper follows the steps: first, each element falls in a cluster of its own, second, clusters are sequentially combined following nearest neighbor clustering, third, two clusters are separated by combining the shortest-distance metric, fourth, calculate the vector distance between all pairs formed by strictly neighboring elements and build a (symmetric) proximity matrix, fifth, in single linkage solution, the distance between two clusters (Q and Z) is the minimum of the distances between all pairs of variable vectors drawn from the two clusters i.e.,

$$D_{x,y} = \min_{i \in Cx, j \in Cy} d(x_i, x_j). \quad (2)$$

The distance measure between a cluster CQ and the new cluster CZ can be obtained by the combinatorial formula:



►Plate 1: Nature of reach-wise variation of sediment grades: **a** coarse (range 1024–2048 mm) flood driven boulder at basin outlet above Patharjhora, **b** gravel channel bed at reach 2 (Turibari) ranging between fine boulder to fine cobble (64–512 mm), **c** bed armor formed by the coating of coarse cobbles to medium pebbles at Manabari (8–128 mm), **d** heterogeneous sediment mostly ranging between coarse to fine cobbles (256–64 mm) at Patharjhora, **e** fining process in sorting comes with monsoon discharge at Odlabari composed of pebble to sand configuration (64–0.063 mm), **f** 1 m² grid sampling frame (field technique), **g** grade wise arrangement of various sediment load to visualize the d_{50} size, **h** channel bar armoring with pebble range size (64–4 mm) after the flood water recession at Odlabari, **i** flood driven sediment load supply of fine cobble to fine pebble size (64–2 mm) to downstream segment near confluence

$$D_{Q,Z} = \frac{1}{2}D_{Q,x} + \frac{1}{2}D_{Q,y} - \frac{1}{2}|D_{Q,x} - D_{Q,y}|, \quad (3)$$

where, D is the shortest distance between clusters X and Y , $d(x, y)$ is the distance between elements in the cluster. Finally, run the process until all clusters are merged in to a single cluster.

Estimation of numbers of clusters

The nearest neighbor method has been considered to know the maximum downstream extension of median grain size movement. In sediment mixing process, heterogeneous sorting is a common phenomenon. But, finding the homogeneity in the downstream sorting process by cluster analysis gives the idea of ongoing fluvial process within the channel. The role of flash flood (Wheatcroft et al. 1997; Ghosh 2019) extended the sediment movement limit within the channel further downstream especially in post-monsoon condition. In this study, the selection of optimal number of clusters is important because the selected clusters will reflect the pattern of sediment size movement downstream. The adjacent pairs in initial clustering will be considered to know the spatio-temporal variability. In the bottom-up approach, we studied the agglomeration coefficients explain within group variance of two clusters combines at each stage. The large change in the agglomeration coefficient value indicates the heterogeneous clusters are started combining. The identification of the ‘knee’ of the agglomeration coefficient indicates the largest magnitude difference between two adjacent points (Chiu et al. 2001; Salvador and Chan 2004; Li et al. 2019; Rodriguez and Laio 2014). Finally, the ANOVA table of K-means clustering confirms the number of selected clusters.

Identity of clusters in bottom-up approach

The geomorphological processes are the end product of complex interaction response of spatial and temporal scale. The structural configuration of various systems (nested,

hierarchic) in process geomorphology gains importance in respect of scale responses (Boer 1992; Zhang et al. 2016, 2017). The present study incorporates the observations on hierarchy of sediment grading process downstream from Haig’s (1987) proposition of hierarchies in dynamics geomorphic systems. The authors have classified the scale response in sediment grading variety downstream in four selected ways: inter-homogeneity (IH-without temporal variability), intra temporal variability (IATV) and inter temporal variability (IETV). At bottom level in hierarchy, the adjacent similar pairs are being considered to check the geomorphological (spatio-temporal integration) scale of response. The responses are frequent at the base of dendrogram (Yim and Ramdeen 2015; Clubb et al. 2019) which considers short distance similarity in the clusters. With increasing distance between clusters, at the top of the dendrogram, we observed only two distinct clusters. Finally, the model aims to understand within and between group temporal variability at the up and down stream segments of the river.

Results and discussion

Results

Descriptive statistics of sediment sampling

In pre-monsoon period, the maximum variance in Normal probability distribution curve (NPDC) with outliers of infrequent deviations has been observed at reach 2 (Fig. 3). In both the seasons, the Shapiro–Wilk test indicates non-normal distribution of frequency except reach 9 in 2016 pre-monsoon and reach 2 and 5 in post-monsoon 2017 ($p < 0.05$). In connection to check the non-normal distribution, the bubble plots (pre- and post-monsoon) exhibit gradual fineness of the sorting process downstream (Figs. 4b, 5b). The heterogeneity of sediment mixing process and multi-modal distribution of samples along a reach both have been presented by 3-D surface plot (Figs. 4c, 5c). The downstream direction is having a tendency of sharp tailing towards right indicates the less probability of observing large boulders (1024–4096 mm) (Plate 1a). There exist a reach-wise random behavior in the peaking tendency of the NPDC (Figs. 4a, 5a). The good Gaussian fit of the NPDC can be observed for the lower reaches. The near outlet reaches indicates more anomaly in the fitting. The skewness and kurtosis values both in pre- and post-monsoon (2015–2017) condition do not exhibit high variation. Negative kurtosis of frequency distribution has been detected for the study reaches means that the outlier character of the data is less extreme. The asymmetry of the FDC is comparatively high near the mountain outlet (Fig. 6a, b). The reach 2 shows sharp tailing near the outlet having the probability of holding maximum

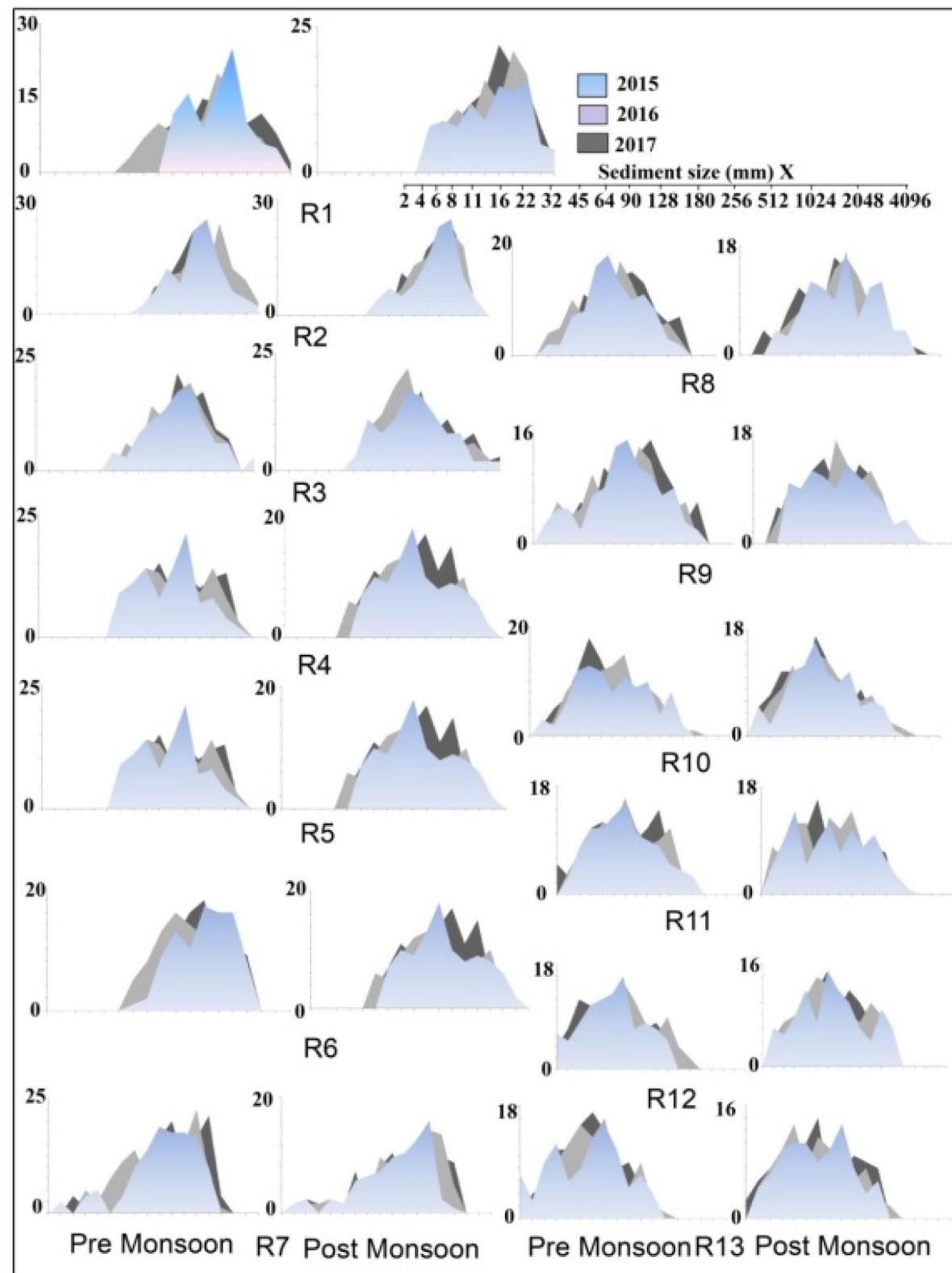


Fig. 3 Reach wise frequency distribution of sediment samples in pre- and post-monsoon conditions

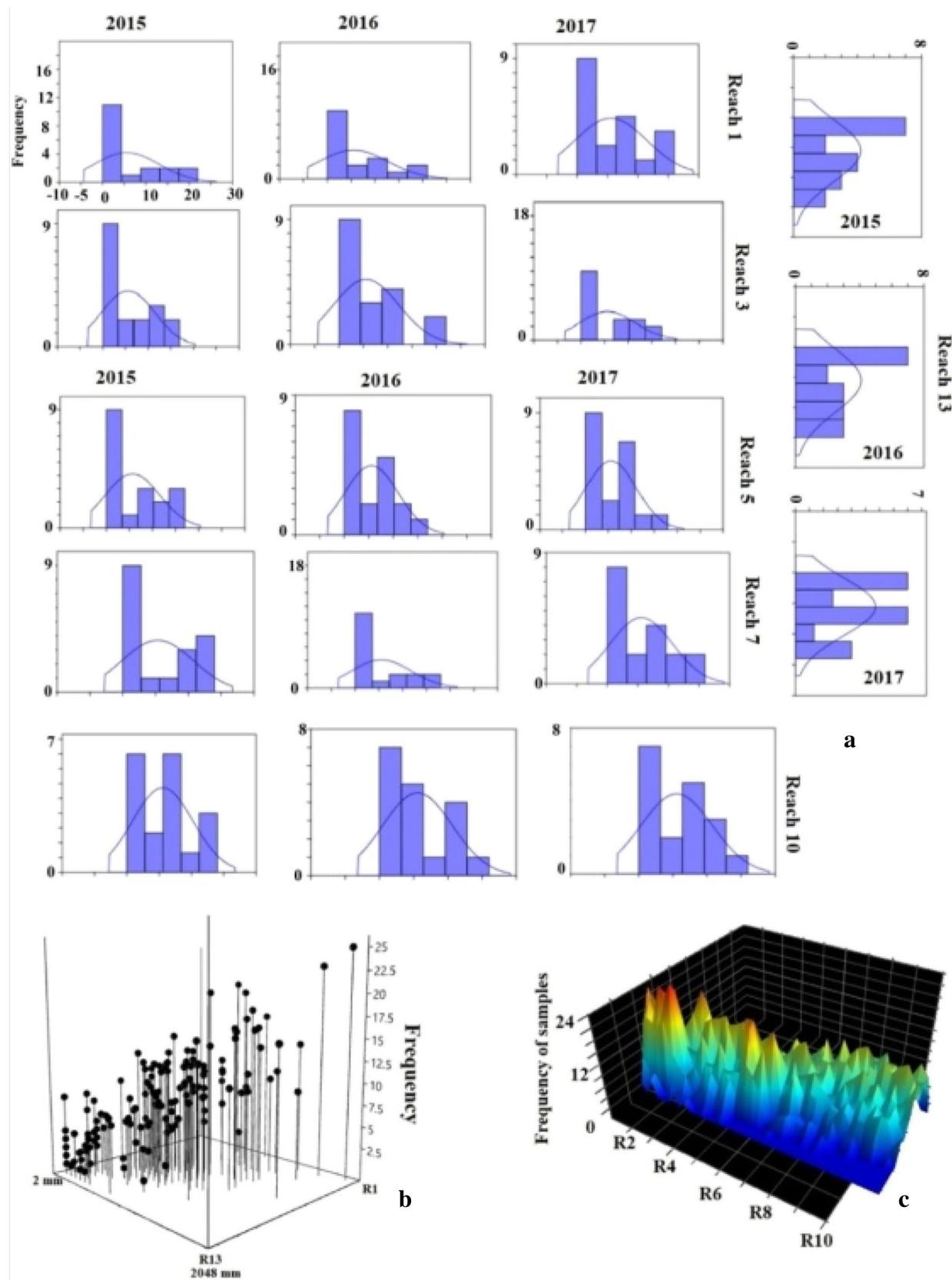


Fig. 4 Distribution of sediment samples in pre-monsoon condition: **a** temporal variation of selected reach wise NPDC, **b** downstream variation of sediment sampling distribution and **c** fluctuating nature of sediment distribution frequency at various grades of Udden-Wentworth scale

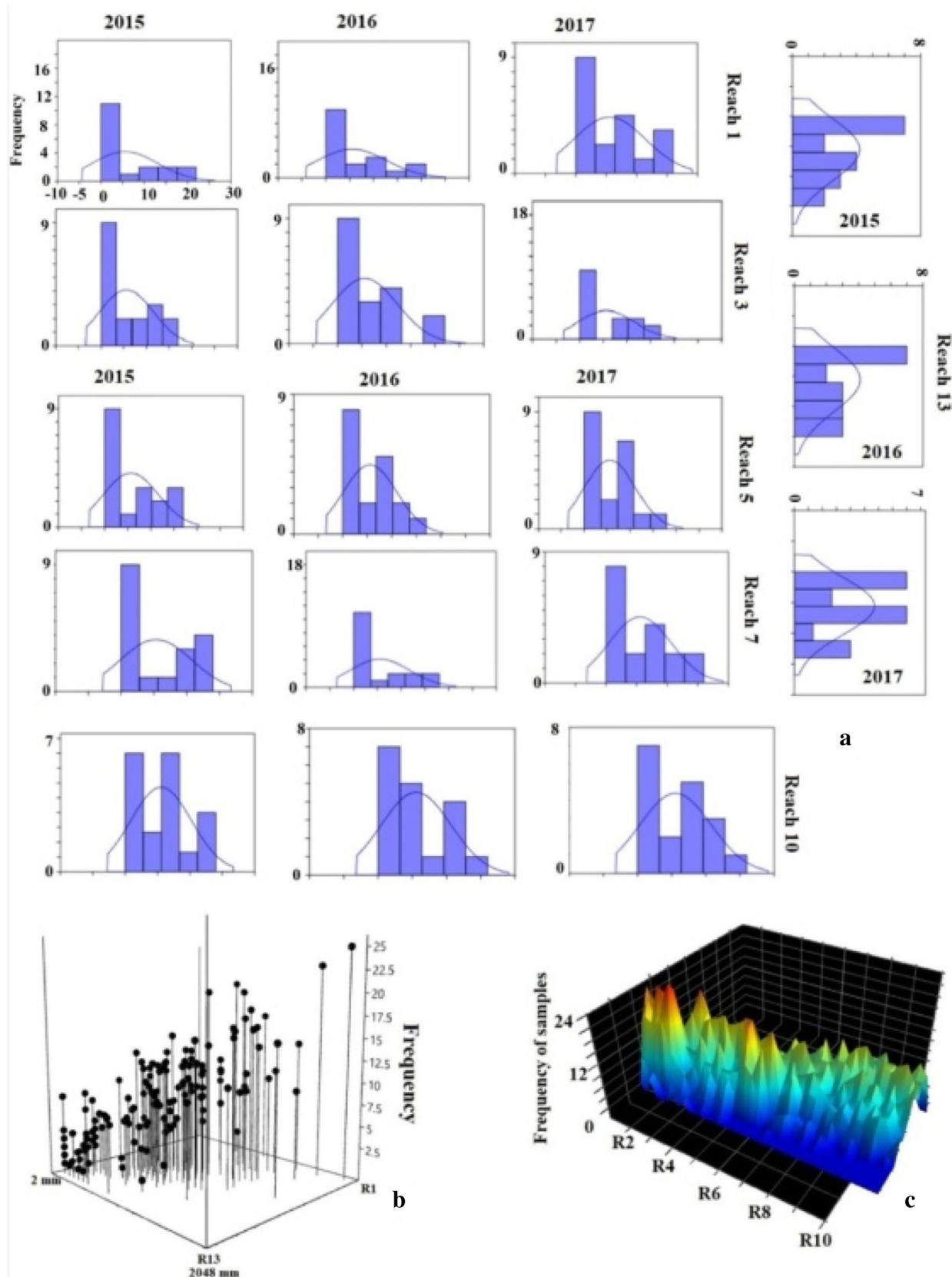


Fig. 5 Distribution of sediment samples in post-monsoon condition: **a** temporal variation of selected reach wise NPDC, **b** downstream variation of sediment sampling distribution and **c** fluctuating nature of sediment distribution frequency at various grades of Udden-Wentworth scale

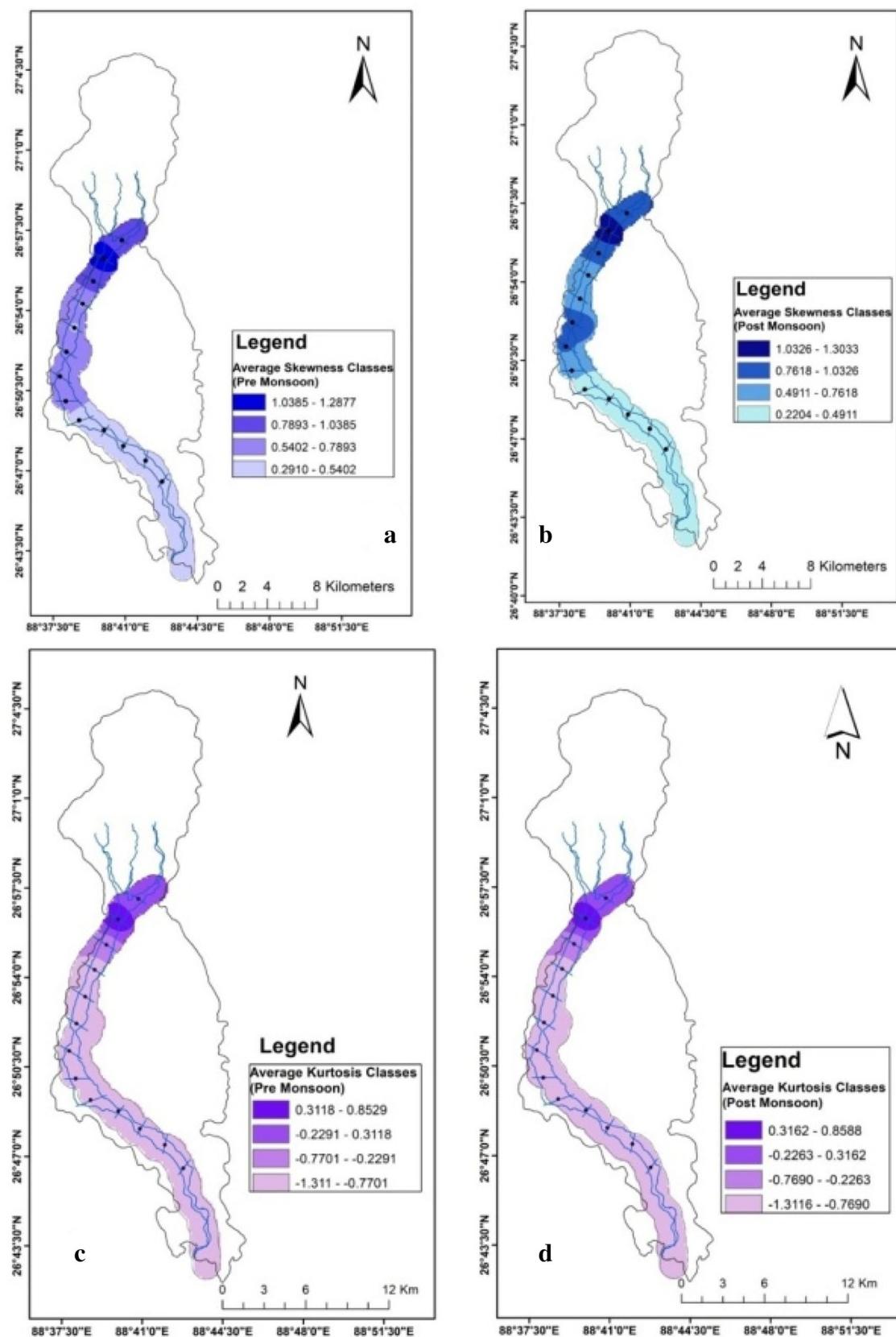


Fig. 6 Spatio-temporal variation of results of skewness and kurtosis values (average): **a** pre-monsoon skewness classes (2015–2017), **b** post-monsoon skewness classes (2015–2017), **c** pre-monsoon kurtosis classes (2015–2017), and **d** post-monsoon kurtosis classes (2015–2017)

Table 1 Descriptive statistics

	2015 (pre)		2016 (pre)		2017 (pre)		2015 (post)		2016 (post)		2017 (post)	
	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min
Kurtosis	–1.766	–0.068	1.003	–0.038	1.649	–0.477	0.034	–0.02	0.735	–0.410	1.725	–0.430
Skewness	1.117	0.203	1.203	0.262	1.555	0.173	1.222	0.223	1.183	0.153	1.578	0.075
Variance	59.908	23.084	50.379	21.794	60.379	22.849	70.970	24.967	67.790	22.026	62.614	22.143

Table 2 One sample *t* test

	<i>t</i>	<i>df</i>	Sig. (2-tailed)	Mean difference
Pre monsoon	3.726	12	.003	989.91780245
Post monsoon	2.868	12	.014	2569.28634320

frequency distribution of 256–512 mm size (Fig. 6c, d). The same reach exhibits less extreme boulder size (outlier) in grading process in 2016 and 2017 pre-monsoon situation. In 2017 post-monsoon, the same reach indicates the existence of flood driven extreme boulder sizes. Near the outlet, the positive kurtosis value (1.725) with positive skewness (1.222) both together indicate maximum sample distribution along the right tail and the frequency distribution peaked at boulder class size (Table 1). The reach-wise variance separately in pre- and post-monsoon is statistically significant (Table 2). But, the comparison of mean difference in pre- and post-monsoon is not statistically different.

Median particle diameter and distribution (pre- and post-monsoon)

On the piedmont, the bed configuration of Chel is highly boulder infested. The average cumulative percentage share

of boulder size (> 256 mm) in the bed remains greater than 86% in pre-monsoon condition (2015–2017). In post-monsoon, it remains greater than 84%. The average of d_{50} size ranges from cobble to boulder both in pre- and post-monsoon condition (Table 3). The pre- and post-monsoon variance in d_{50} size is not statistically different. The range of 90–512 mm size holds the possibility of getting median size range in every study reaches (Table 3). The grade of d_{50} load is comparatively higher in post-monsoon near the outlet.

Discussion

Understanding reach-wise surface sediment heterogeneity (SSH) (> 2 mm)

The regulation of flood discharge sets various thresholds in respect to sediment supply limitation. The disappearing fan surface at Patharjhora supports maximum range of heterogeneity of surface sediment load controlled by various intensity of flood discharge and seepage action of channel. This segment focuses on the results of pebble count method. At this reach, the range of sediment sampling grades increases in post-monsoon period (128–1024 mm) compared to pre-monsoon period (64–256 mm) (Fig. 7). It indicates the role of occasional floods (Ghosh 2019) and its varying power of generating surface movement of loads. From outlet

Table 3 Downstream variation of median (d_{50}) particle size (mm)

Reaches	Pre monsoon	Pre monsoon	Pre monsoon	Post monsoon	Post monsoon	Post monsoon
	2015	2016	2017	2015	2016	2017
1	348.62	374.32	435.62	451.23	445.37	512.76
2	423.43	412.73	401.23	501.34	489.15	554.21
3	327.36	276.48	344.58	337.78	346.73	367.29
4	388.52	359.61	347.17	376.32	259.75	401.32
5	321.23	274.82	263.32	301.12	345.23	189.56
6	348.83	289.53	236.73	233.45	178.23	256.63
7	262.32	177.22	184.75	267.21	245.71	305.82
8	275.21	269.71	256.14	165.23	165.37	163.69
9	165.92	154.63	201.32	227.74	201.54	302.41
10	226.13	163.21	225.31	79.54	123.49	278.94
11	104.32	78.86	111.55	101.32	158.79	154.32
12	88.48	102.53	124.28	84.12	78.84	101.82
13	98.23	75.89	84.53	58.27	112.57	57.22

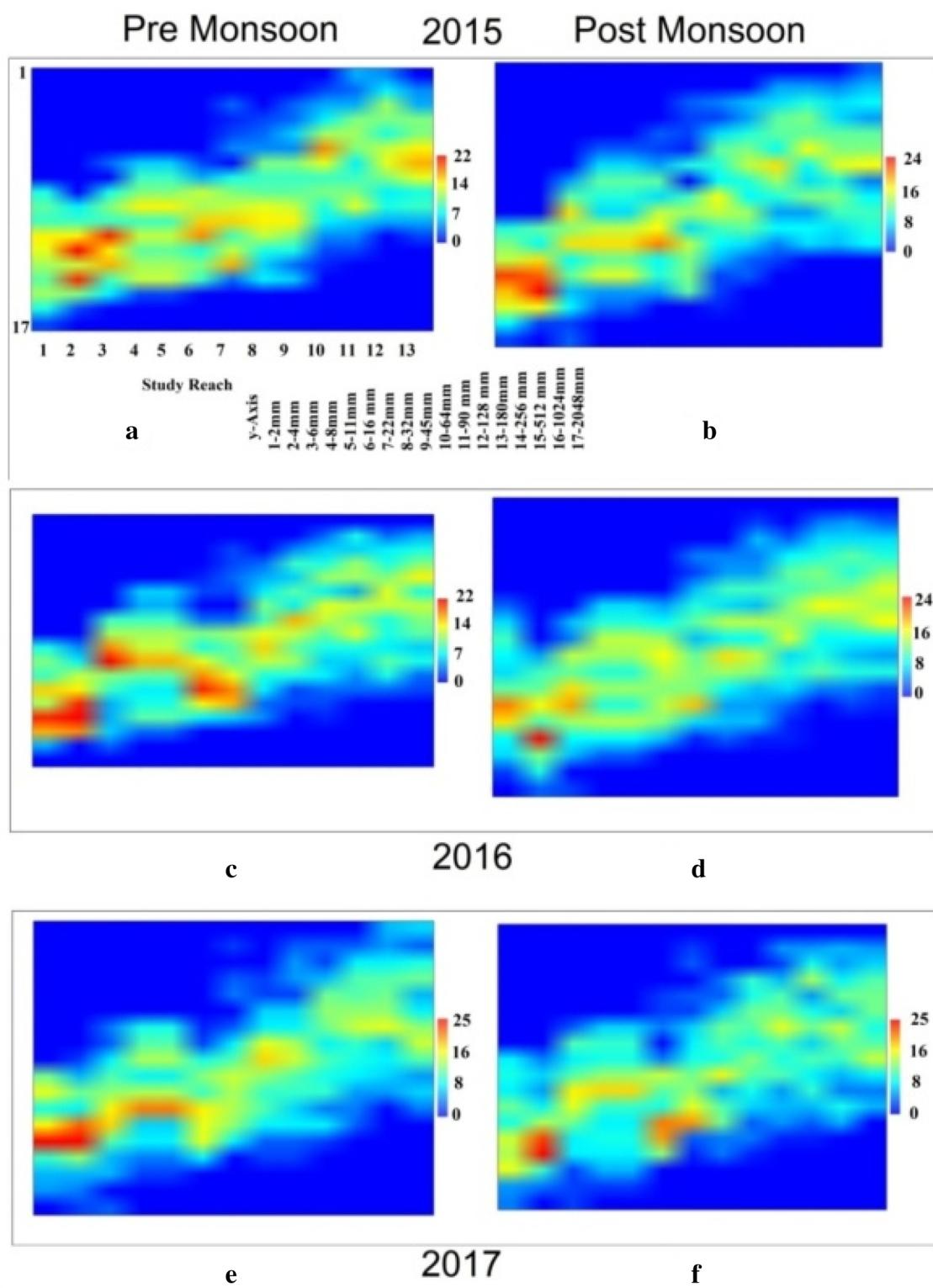


Fig. 7 Surface matrix plot (SMP) of sampling distribution: **a** pre-monsoon (2015), **b** post-monsoon (2015), **c** pre-monsoon (2016), **d** post-monsoon (2016), **e** pre-monsoon (2017) and **f** post-monsoon (2017)

to 4 km downstream, the presence of isolated boulders (2048–4096 mm and greater than 4096 mm) is noticeable.

In post-monsoon period, the seasonal effect of flood frequencies drag the small boulders (256–512 mm) up to Manabari reach (6 km downstream) (Plate 1b). Up to the extension of active fan surface (8 km from outlet), the post-monsoon composition of bed indicates the presence of small cobble to medium boulders (130–832 mm) connecting the extreme limit of boulder supply in the channel (Plate 1c). In October, the appearance of extended pebble bars (Plate 1h) (20–60 mm) make the water flow to some extent rampant along the edges of the medial gravel bars. The post-monsoon range of sediment grades (10–512 mm) slightly vary from pre-monsoon grades (22–512 mm). At Apalchad reach (14 km downstream) (Plate 1i), the channel gets diverted into two branches along a mid-channel bar. The range of sediment grade is maximum here (4–250 mm) varies from coarse cobble to fine pebble. The formation of shoals is initiated from this reach. After the flood of 2017, the surface composition of the shoals varies from very fine to coarse pebble (4–32 mm).

Understanding the sorting process from frequency distribution

The shape of the frequency distribution curve has the capability to show the gradual sorting process within the channel. The analysis technique is considered to understand the deviation of the frequency distribution curve from normality. The channel sets the supply limits of each size class. The overall frequency distribution from reach 1–13 both in pre- and post-monsoon has been considered in the discussion. The Shapiro–Wilk's test with insignificant test value ($p < 0.05$) proves the non-normal distribution of the sediment samples. The NPDC with standard deviation ranges indicates the similarity of the sampling distribution with the theoretical standard normal probability curve. The reach-wise frequency distribution shows the right tailing of the curve. The positive skewness values with gradual flattening

towards right tail indicates the probability of rare existence of high sediment grades like medium to large boulders (> 1024 mm) and fine boulders (> 512 mm) up to reach 3 and below that respectively. The peak forms within a range of coarse to fine cobble (128–64 mm) within the reaches. The probability theory also measures the ‘tailedness’ of the frequency distribution by kurtosis. The kurtosis of any univariate normal distribution approximates to 3. In pre-monsoon, the positive kurtosis values are observed at reach 1 and 2. It indicates the outlier character of the distribution is extreme. So, the probability of getting coarse boulders (> 1024 mm) near the outlet is high. In post-monsoon, the runoff drives the coarse boulders further downstream. In 2017 post-monsoon condition, the high variation in the sediment distribution process can be observed. Both in 2017 and 2016, the sharp decline of positive kurtosis values up to reach 3 indicates the outlier character of the distribution is more extreme near the mountain outlet reaches than the downstream reaches.

Downstream variation of d_{50} particle size

The median particle diameter (d_{50}) represents 50% of the particle size falls less than d_{50} diameter. In particle size distribution, this median value is important to determine the standard critical share of stress needed to start the incipient motion, the amount of terminal velocity and stream power for sediment mobility. The range of pre- and post-monsoon variation in d_{50} size has been calculated to vary from 412.46 to 86.21 mm and 514.90 mm to 76.02 mm respectively. In 2016 (pre- and post-monsoon), the maximum d_{50} variation in downstream direction has been observed (412.46–86.21 mm and 514.90–76.02 mm) (Table 3). At 2 km downstream, the range of post-monsoon d_{50} size is observed to be the highest (489.15–554.21 mm). The lowest range of d_{50} size variation is observed in pre-monsoon at 26 km downstream i.e. 75.89–98.23 mm. In lower reaches, the pre- and post-monsoon standard deviation (SD) is maximum due to the variable nature of sediment threshold limits along the channel.

Table 4 ANOVA results in selecting clusters for median sediment particle size (> 2 mm)

	Cluster		Error		<i>F</i>	Sig.
	Mean square	<i>df</i>	Mean square	<i>df</i>		
Pre-monsoon (2015)	116,923.260	1	3702.095	11	31.583	.000
Pre-monsoon (2016)	125,473.863	1	2775.201	11	45.213	.000
Post-monsoon (2017)	94,576.183	1	4631.004	11	20.422	.001
Post-monsoon (2015)	170,242.470	1	7531.202	11	22.605	.001
Post-monsoon (2016)	117,927.438	1	7467.967	11	15.791	.002
Pre-monsoon (2017)	183,656.925	1	8040.937	11	22.840	.001

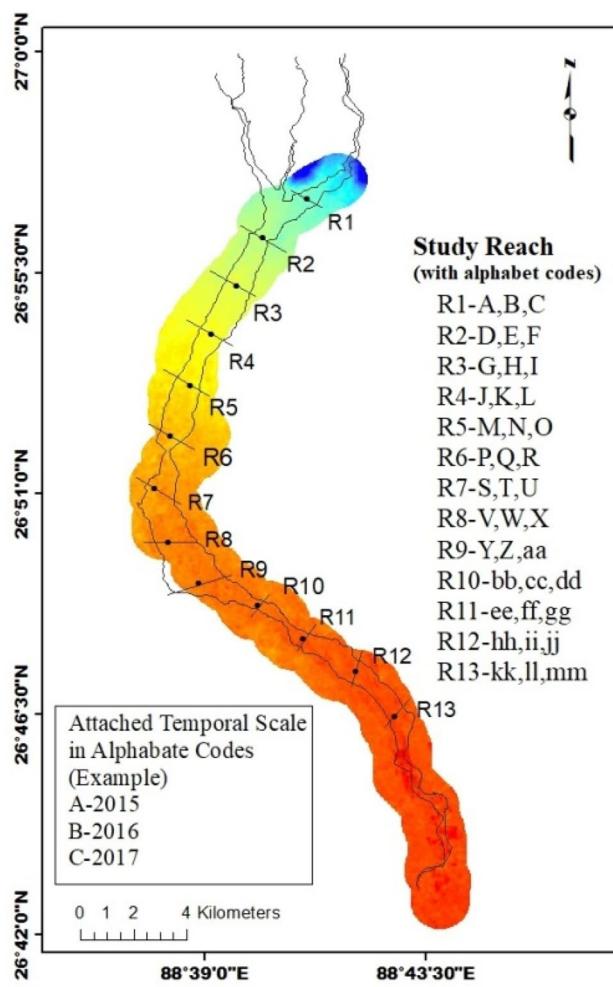


Fig. 8 Sampling reaches with alphabet codes of HCA

The range of d_{50} size distribution is observed maximum in 2017 post-monsoon. The F -test for two sample variance (pre-monsoon-989.917 and post-monsoon-2569.286) is calculated to be significantly different from each other (Table 4). This result promotes the final question that how we can capture the similar variances in form of clusters to know the sediment mixing process and sediment supply limits in pre- and post-monsoon condition.

Understanding the sediment mixing from clustering

The nature of sediment distribution is highly erratic and varies significantly. It depends on chance that different reaches may have similarity or dissimilarity in clustering process. The concept develops on the limits of sediment supply to downstream. The sediment load dispersal generally follows the process of sorting at downstream. But, recurring occasional floods with varying intensity and rainstorm events are frequently changing the sediment supply process in the channel. So, the sediment mixing disrupts the sorting process simultaneously.

Finding sediment supply limit of d_{50} size

The study reaches are being coded with alphabets to know the spatial similarity in d_{50} size mixing (Fig. 8). In pre-monsoon the ranges of d_{50} size varies from fine boulder to fine cobble (Fig. 9a). The lower reaches (11–13) form a small range cluster and the second cluster incorporates reach 1–10. The upper reaches (1–10) comprise median particle size from medium boulder to coarse cobble. The lower reaches (11–13) vary within fine cobble to very coarse pebble. In pre-monsoon, the supply limit of boulders is limited up to reach 10. The range gets widened in post-monsoon and varies from medium boulder to very coarse pebble. Up

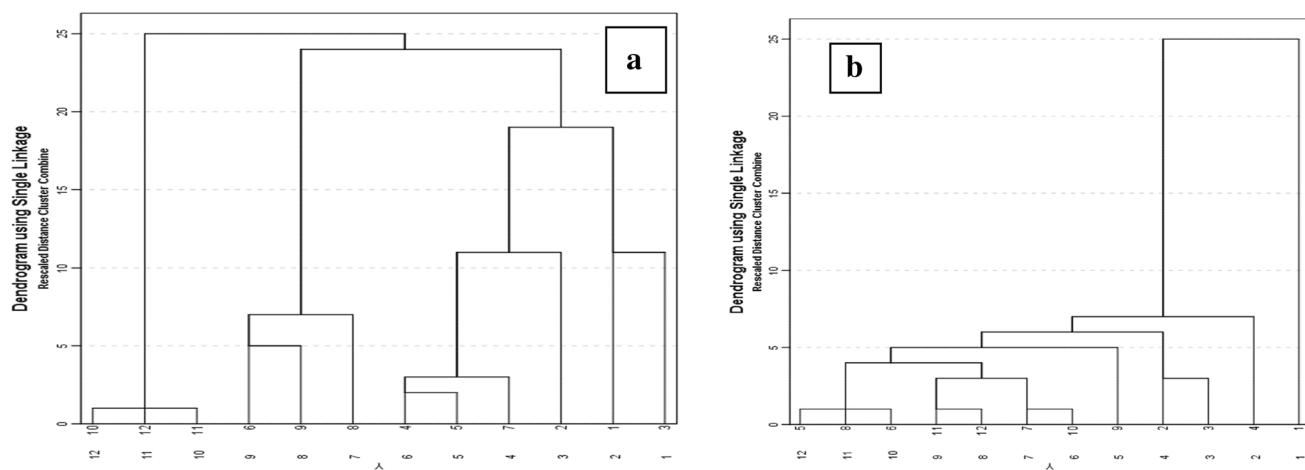


Fig. 9 Dendrogram (simple linkage) of d_{50} particle size: **a** pre-monsoon, **b** post-monsoon

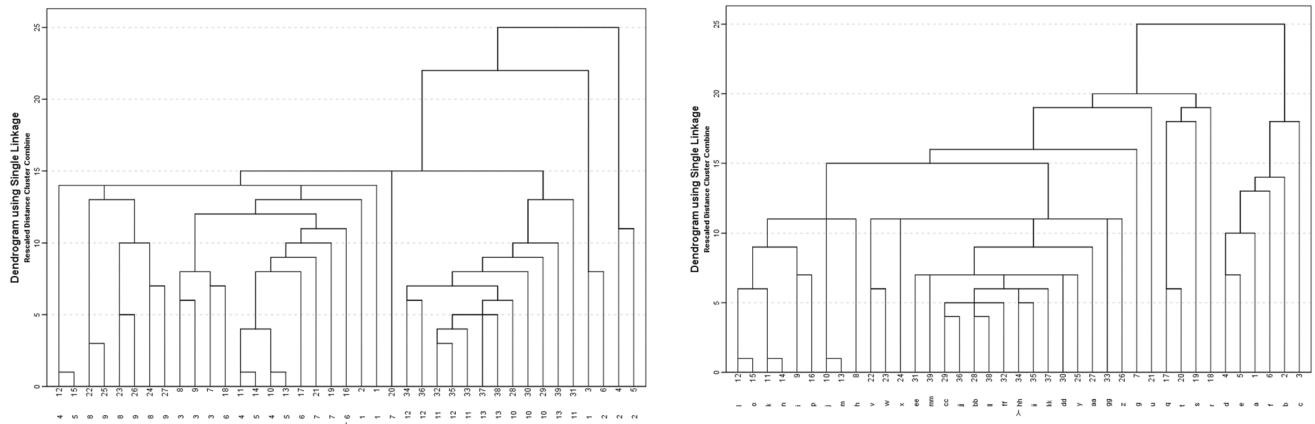


Fig. 10 Dendrogram of overall size clusters: **a** pre-monsoon, **b** post-monsoon

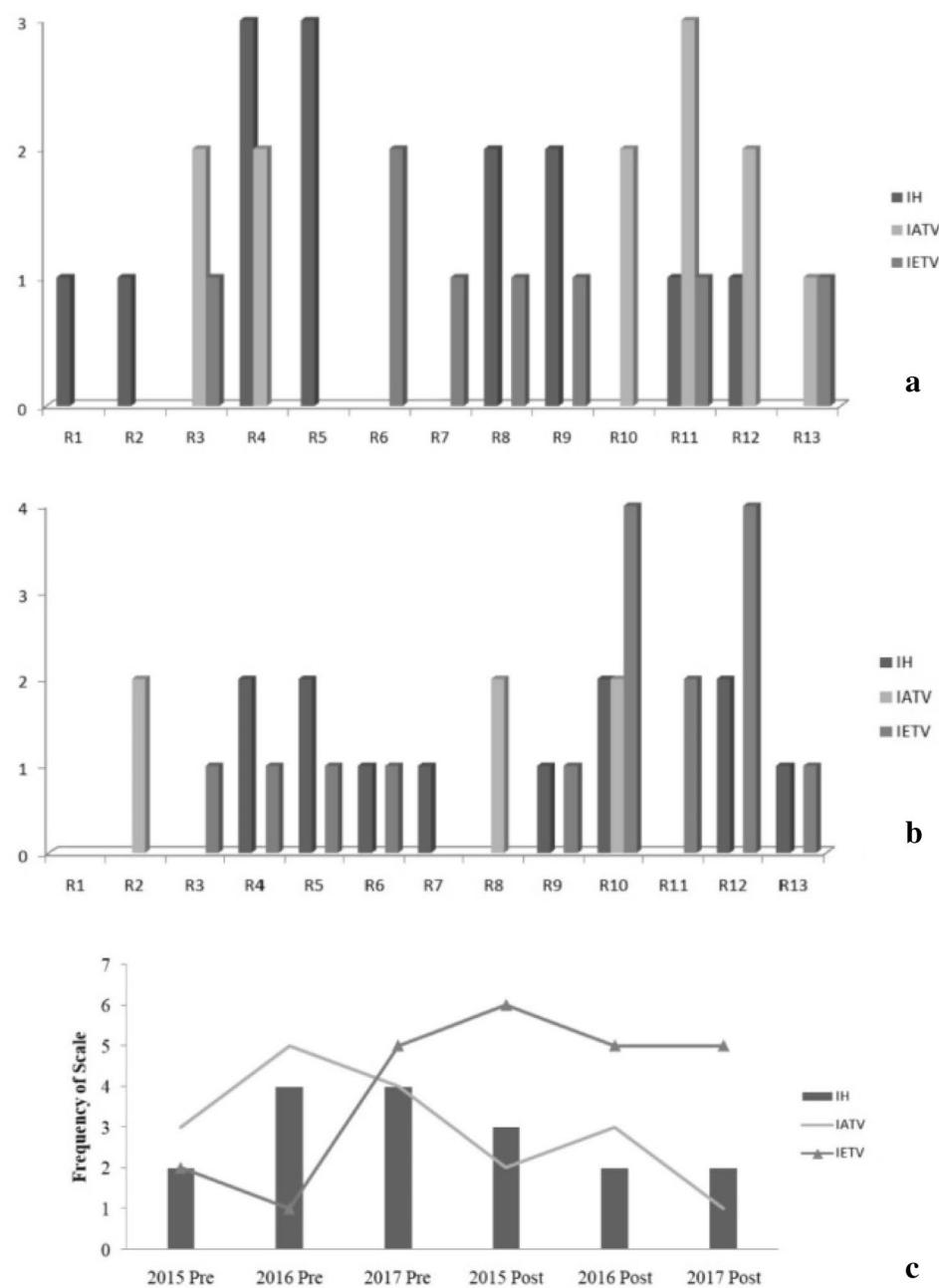
to reach 2, the range varies from medium to fine boulder (Fig. 9b). The probability of getting large sediment size is high at these sites which also connects the positive value of kurtosis (Fig. 6d). The fineness of sediment sorting increases in post-monsoon condition.

Understanding sediment mixing process from overall size class clustering

In pre-monsoon, the dendrogram indicates similarities among reach 1–3 (cluster 1), reach 4–9 (cluster 2) and reach 10–13 (cluster 3) (Fig. 10a). In 2017, with exception reach 6 shows similarity in sediment mixing process. In cluster 1, 2 and 3 the dominant d_{50} size varies from fine boulder, coarse to fine cobble and coarse cobble to fine pebble respectively. In post-monsoon, the dendrogram indicates similarities among reach 1–2 (cluster 1), reach 3–7 (cluster 2) and reach 8–13 (cluster 3) (Fig. 10b). In 2017, the reach 8 shows similarity with cluster 2 in the mixing process. The dominant d_{50} size varies from fine boulder, fine to coarse cobble and coarse cobble to very coarse pebble respectively in the selected clusters. Finally, the intra and inter homogeneity and variability has been calculated from the pairs of initial clusters. The technique addresses the spatio-temporal variability in the sediment mixing process downstream (Fig. 11a–c). In cluster analysis, the IATV and IETV responses are considered to know the within and across variation of the clusters from initial level. But, the analysis produces IH response to the between group homogeneity (across) which is the rule base in clustering process. The aim is to check the between

and within group homogeneity of the reaches and their readjustment with time scale. The between group homogeneity (IH) is observed between the intermediate reaches (4, 5, 6 and 7) and lower reaches (9, 13) in pre-monsoon. In post-monsoon, the homogeneity has been observed between the upper reaches (1, 2, 4 and 5) and lower reaches (8, 9, 11 and 12). In the lower reaches, maximum sediment mixing is observed due to the role of flood discharge. Here, the process of sediment mixing has been observed at a particular time frame. IATV response captures the within group temporal difference. The reaches 8 and 10 have similarity in sediment mixing process within the time frame (2015–2017) in pre-monsoon. In post-monsoon, both the upper (3, 4) and lower (10–12) reaches have the similarity. The most comprehensive both spatio-temporal responses can be captured by the IETV response (between group differences). These differences make each sampling site unique. The uniqueness of each sites are being conceptualized on the supply limitation of d_{50} size previously. The IETV responses have been captured covering maximum number of reaches (3–6 and 9–13) in pre-monsoon. The sediment rolling process on the bed gets more time to sparsely distribute the bed load in pre-monsoon. In post-monsoon, the lower reaches (8–13) indicate more comprehensive sediment mixing process due to the flood water thrust, breaks the sediment supply limit of d_{50} particle size. The error graphs (pre- and post-monsoon) provide the mark of selecting reach 8 and 7 as the mixing zone between boulder and cobble d_{50} size (Fig. 12a, b). The role of flood discharge invariably contributes to bring similarity among furthest reaches by chance.

Fig. 11 Variation in scale of response from cluster analysis: **a** spatial scale response in pre-monsoon, **b** spatial scale response in post-monsoon and **c** variation of temporal scale response



Conclusion

The study produces two aspects like finding the reach wise differences in the sediment mixing process downstream and the sediment supply limits of median particle size. In pre-monsoon, reach 8 is brought out as the supply limit of boulder size (> 256 mm) from 2015 to 2017. In post-monsoon, reach 7 is observed as the supply limit of boulder size. Below the reach 7 the range of sediment sorting expanses up to a smaller pebble size (< 64 mm) in 2015 and 2017 flood years. The reaches near the outlet are having representative d_{50} size as fine boulders (256–512 mm). The reaches near

the confluence are having heterogeneous representation of d_{50} size ranging between fine cobbles to very coarse pebble (32–128 mm). The confirmative clustering defines the dendrogram rightly. The between group difference reflects in the downstream difference in sediment mixing process in spatial scale. But, the IETV response includes time difference (2015–2017) which reflects the variation in stream competency. In Post monsoon, the downstream dispersal of fine sediment grades is more favorable for coarse cobble to very coarse pebble grades. The boulder sizes get less chance of downstream movement (up to reach 2). The intermediate size (coarse to fine cobble) gets more chance of its dispersal

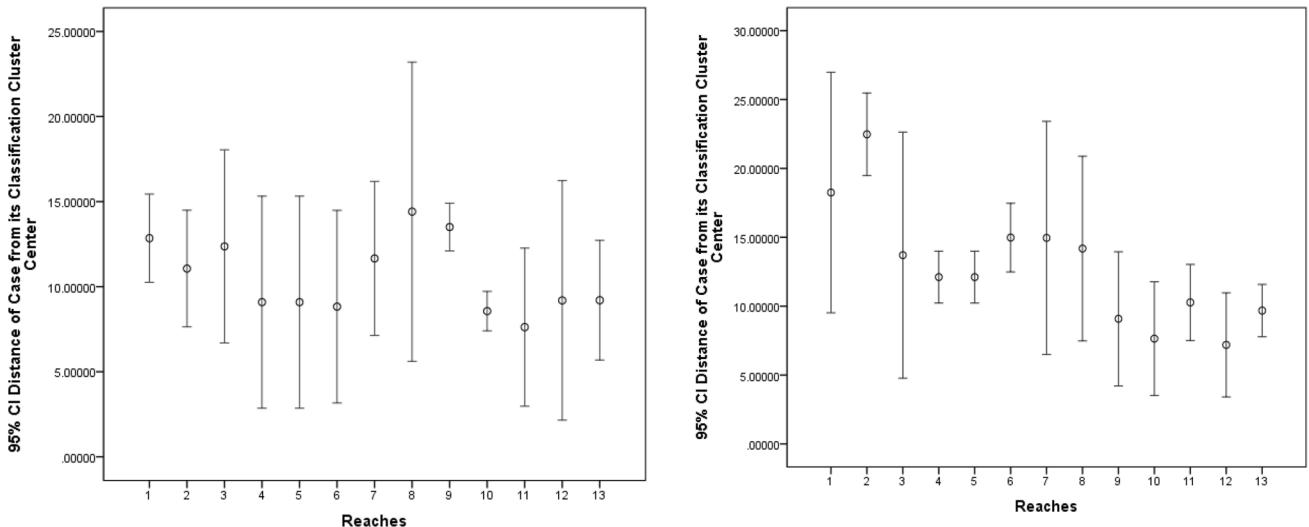


Fig. 12 Error graph of reach-wise frequency distribution of overall size class: **a** pre-monsoon, **b** post-monsoon

and mixing to downstream in pre-monsoon. The clusters of finding overall similarities in the mixing process among the reaches come up with three responses (IH, IATV and IETV). It comprises finding all possibilities of sediment mixing process. The bottom up approach incorporates in finding the pairs of similar reaches as clusters and ends with capturing the reaches with variability in sediment mixing in a temporal scale.

Acknowledgements The researchers are indebted and grateful to the Department of Geography and Applied Geography, University of North Bengal especially to Dr. Rupak Kumar Pal to accomplish the whole work. We received valuable suggestions from Mr. Mihir Kumar Roy, B. Tech (H), IIT Bhubaneswar, Dr. Dipayan Datta, BDS and MDS and Dr. Diyali Chattaraj for framing the conceptual background of the paper. Finally, the urge of writing good scientific paper on geomorphology comes from the courageous notes of Prof. Takashi Oguchi and Prof. Sunil Kumar Dey in Young Geomorphologist workshop in Calcutta University in spring, 2020.

References

- Blair CT, McPherson JG (1999) Grain size and textural classification of coarse sedimentary particles. *J Sediment Res* 69(1):6–19. <https://doi.org/10.2110/jsr.69.6>
- Blott SJ, Pye K (2012) Particle size scales and classification of sediment types based on particle size distributions: review and recommended procedures. *Sedimentology* 59(7):2071–2096. <https://doi.org/10.1111/j.1365-3091.2012.01335.x>
- Boer HD (1992) Hierarchies and spatial scale in process geomorphology: a review. *Geomorphology* 4:303–318
- Boulay S, Colin C, Trentesaux A, Pluquet F, Bertaux J, Blamart D, Wang P (2003) Mineralogy and sedimentology of Pleistocene sediment in the South China Sea (ODP Site 1144). *Proc ODP Sci Results* 184(211):1–21. <https://doi.org/10.1016/j.yqres.2007.03.004>
- Brasington J, Langham J, Rumsby B (2003) Methodological sensitivity of morphometric estimates of coarse fluvial sediment transport. *Geomorphology* 53(3–4):299–316. [https://doi.org/10.1016/S0169-555X\(02\)00320-3](https://doi.org/10.1016/S0169-555X(02)00320-3)
- Bunte K, Abt SR (2001) Sampling surface and subsurface particle-size distributions in wadable gravel-and cobble-bed streams for analyses in sediment transport, hydraulics, and streambed monitoring. US Department of Agriculture, Forest Service, Rocky Mountain Research Station. <https://doi.org/10.2737/RMRS-GTR-74>
- Chakraborty P, Nag S (2015) Rivers of West Bengal: changing scenario, geoinformatics and remote sensing cell. West Bengal State Council of Science and Technology, Department of Science & Technology, Govt. of WB, pp 9–13
- Chiu T, Fang D, Chen J, Wang Y, Jeris C (2001) A robust and scalable clustering algorithm for mixed type attributes in large database environment. In: Proceedings of the seventh ACM SIGKDD international conference on knowledge discovery and data mining, San Francisco, CA, pp 263–268. <https://doi.org/10.1145/502512.502549>
- Church M (2010) The trajectory of geomorphology. *Prog Phys Geogr* 34(3):265–286
- Clark MW (1976) Some methods for statistical analysis of multimodal distributions and their application to grain-size data. *J Int Assoc Math Geol* 8(3):267–282. <https://doi.org/10.1007/BF01029273>
- Clubb F, Brookhaven B, Rheinwald A (2019) Clustering river profiles to classify geomorphic domains. *J Geophys Res Earth Surf* 124(6):1417–1435. <https://doi.org/10.1029/2019JF005025>
- Constantine JA, Dunne T, Ahmed J (2014) Sediment supply as a driver of river meandering and floodplain evolution in the Amazon Basin. *Nat Geosci* 7(2):899–903. <https://doi.org/10.1002/esp.4048>
- Corenblit D, Neil SD, Johannes S, Martin RG, Gudrun B (2015) Considering river structure and stability in the light of evolution: feedbacks between riparian vegetation and hydrogeomorphology. *Earth Surf Proc Land* 40(02):189–207. <https://doi.org/10.1002/esp.3643>
- Dawyndt P, De Meyer H, De Baets B (2006) UPGMA clustering revisited: a weight-driven approach to transitive approximation. *Int J Approx Reason* 42(3):174–191. <https://doi.org/10.1016/j.ijar.2005.11.001>
- De Carvalho FDA, Lechevallier Y (2009) Partitional clustering algorithms for symbolic interval data based on single

- adaptive distances. *Pattern Recognit* 42(7):1223–1236. <https://doi.org/10.1016/j.patcog.2008.11.016>
- Doeglas DJ (1968) Grain size indices, classification and environment. *Sedimentology* 10:90–92
- Donato SV, Reinhardt EG, Boyce JI, Pilarczyk JE, Jupp BP (2009) Particle-size distribution of inferred tsunami deposits in Sur Lagoon. *Sultanate Oman Mar Geol* 257(1–4):54–64. <https://doi.org/10.1016/j.margeo.2008.10.012>
- Dutta I, Das A (2019) Modeling dynamics of peri-urban interface based on principal component analysis (PCA) and cluster analysis (CA): a study of English Bazar Urban Agglomeration, West Bengal. *Model Earth Syst Environ* 5(2):613–626. <https://doi.org/10.1007/s40808-018-0554-6>
- Dutta D, Sil J, Dutta P (2019) Automatic clustering by multi-objective genetic algorithm with numeric and categorical features. *Expert Syst Appl* 137:357–379. <https://doi.org/10.1016/j.eswa.2019.06.056>
- Everitt BS, Landau S, Leese M, Stahl D (2011) Hierarchical clustering. *Clust Anal* 5:71–110
- Folk RL (1954) The distinction between grain size and mineral composition in sedimentary rock nomenclature. *J Geol* 62:344–355
- Fournier J, Gallon RK, Paris R (2014) G2Sd: a new R package for the statistical analysis of unconsolidated sediments. *Géomorphologie Relief Process Environ* 20(1):73–78. <https://doi.org/10.4000/geomorphologie.10513>
- Gayer E, Mukhopadhyay S, Meade BJ (2008) Spatial variability of erosion rates inferred from the frequency distribution of cosmogenic ^{3}He in olivines from Hawaiian river sediments. *Earth Planet Sci Lett* 266(3–4):303–315. <https://doi.org/10.1016/j.epsl.2007.11.019>
- Ghosh D (2019) Determining process of occasional flooding from channel hydrological characteristics of Chel Basin, North Bengal. *Spat Inf Res* 28:313–326. <https://doi.org/10.1007/s4132-019-00290-0>
- Haig BD (1987) Scientific problems and conduct of research. *Educ Philos Theory* 19:22–32
- Han J, Lee JG, Kamber M (2009) An overview of clustering methods in geographic data analysis. *Geogr Data Min Knowl Discov* 2:149–170
- Hayashi H, Shimatani Y, Shigematsu K (2012) A study of seed sediment dispersal by flood flow in an artificially restored floodplain. *Landsc Ecol Eng* 8:129–130. <https://doi.org/10.1007/S1135-5-011-0154-3>
- Hickin EJ (1995) River geomorphology. Wiley Publication, New York, pp 70–106
- Jerosch K (2013) Geostatistical mapping and spatial variability of surficial sediment types on the Beaufort Shelf based on grain size data. *J Mar Syst* 127:5–13. <https://doi.org/10.1016/j.jmarsys.2012.02.013>
- Julian PV (1998) Erosion and sedimentation. Cambridge University Press, Cambridge, pp 4–81
- Karen BG, Michal T, Emily DW (2015) Co-evolution of riparian vegetation and channel dynamics in an aggrading braided river system, Mount Pinatubo, Philippines. *Earth Surf Process Landf* 40(08):1101–1115. <https://doi.org/10.1002/esp.3699>
- Kellerhals R, Bray DI (1971) Sampling procedures for coarse fluvial sediments. *J Hydraul Div Am Soc Civ Eng* 97(HY8):1165–1179
- Kleinmans M (2002) Sorting out sand and gravel: Sediment transport and deposition in sand-gravel bed rivers. *Neth Geogr Stud* 293:25–40
- Kondlof GM, Li S (1992) The pebble count technique for quantifying surface bed material size in instream flow studies. *Rivers* 3(2):80–87
- Lamb DS, Downs JA, Lee C (2016) The network K-function in context: examining the effects of network structure on the network K-function. *Trans GIS* 20(3):448–460. <https://doi.org/10.1111/tgis.12157>
- Lamb DS, Downs J, Reader S (2020) Space-time hierarchical clustering for identifying clusters in spatiotemporal point data. *ISPRS Int J Geo-Inf* 9(2):85. <https://doi.org/10.3390/ijgi9020085>
- Langfelder P, Zhang B, Horvath S (2008) Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics* 24(5):719–720. <https://doi.org/10.1093/bioinformatics/btm563>
- Leys J, McTainsh G, Koen T, Mooney B, Strong C (2005) Testing a statistical curve-fitting procedure for quantifying sediment populations within multi-modal particle-size distributions. *Earth Surf Process Landf J Br Geomorphol Res Group* 30(5):579–590. <https://doi.org/10.1002/esp.1159>
- Li Z, Zhang Y, Xu H, Li K, Dubovik O, Goloub P (2019) The fundamental aerosol models over China region: a cluster analysis of the ground-based remote sensing measurements of total columnar atmosphere. *Geophys Res Lett* 46(9):4924–4932. <https://doi.org/10.1029/2019GL082056>
- Liu X, Zhou S, Wang Y, Li M, Dou Y, Zhu E, Li H (2017) Optimal neighborhood kernel clustering with multiple kernels. In: AAAI, pp 2266–2272
- López G (2017) Encyclopedia of geoarchaeology. In: Gilbert AS (ed) *Grain size analysis*. Springer, Israel, pp 341–348. <https://doi.org/10.1007/978-1-4020-4409-0>
- Lu W, Atkinson DE, Newlands NK (2017) ENSO climate risk: predicting crop yield variability and coherence using cluster-based PCA. *Model Earth Syst Environ* 3(4):1343–1359. <https://doi.org/10.1007/s40808-017-0382-0>
- Luo C, Zheng Z, Zou H, Pan A, Fang G, Bai J, Li J, Yang M (2013) Palaeo environmental significance of grain size distribution of river flood deposits: a study of the archeological sites of the Apengjiang River Drainage, upper Yangtze region, Chongqing China. *J Archol Sci* 40(2):827–832. <https://doi.org/10.1007/s12040-018-1030-4>
- Ma H, Nittrouer JA, Naito K, Fu X, Zhang Y, Moodie AJ, Wang Y, Wu B, Parker G (2017) The exceptional sediment load of fine grained dispersal systems: examples of the Yellow River, China. *Sedimentology* 3(5):1–7. <https://doi.org/10.1126/sciadv.1603114>
- Mandal S, Sarkar S (2016) Overprint of Neotectonism along with the course of River Chel, North Bengal, India. *J Palaeogeogr* 5(3):222–223. <https://doi.org/10.1016/j.jop.2016.05.004>
- Manson SM (2008) Does scale exist? An epistemological scale continuum for complex human-environment systems. *Geoforum* 39(2):776–788. <https://doi.org/10.1016/j.geoforum.2006.09.010>
- Nelson PA, Bellugi D, Dietrich WE (2014) Delineation of river bed-surface patches by clustering high-resolution spatial grain size data. *Geomorphology* 205:102–119. <https://doi.org/10.1016/j.geomorph.2012.06.008>
- Okabe A, Yamada I (2001) The K-function method on a network and its computational implementation. *Geogr Anal* 33(3):271–290. <https://doi.org/10.1111/j.1538-4632.2001.tb00448.x>
- Ordóñez C, Ruiz-Barzola O, Sierra C (2016) Sediment particle size distributions apportionment by means of functional cluster analysis (FCA). *CATENA* 137:31–36. <https://doi.org/10.1016/j.catena.2015.09.006>
- Passega R (1964) Grain size characteristics by CM pattern as a tool. *J Sediment Petrol* 34:233–247
- Paterson GA, Heslop D (2015) New methods for unmixing sediment grain size data. *Geochem Geophys Geosyst* 16(12):4494–4506. <https://doi.org/10.1002/2015GC006070>
- Poppe LJ, Eliason AH, Hastings ME (2004) A Visual Basic program to generate sediment grain-size statistics and to extrapolate particle distributions. *Comput Geosci* 30(7):791–795. <https://doi.org/10.1016/j.cageo.2004.05.005>
- Qiu Z, Sun D, Hu C, Wang S, Zheng L, Huan Y, Peng T (2016) Variability of particle size distributions in the Bohai Sea and the Yellow Sea. *Remote Sens* 8(949):2–19. <https://doi.org/10.3390/rs8110949>

- Rinaldo A (1991) Geomorphological dispersion. *Water Resour Res* 27(4):513–525
- Rodriguez A, Laio A (2014) Clustering by fast search and find of density peaks. *Science* 344(6191):1492–1496. <https://doi.org/10.1126/science.1242072>
- Salvador S, Chan P (2004) Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms. In: 16th IEEE international conference on tools with artificial intelligence, IEEE, pp 576–584. <https://doi.org/10.1109/ICTAI.2004.50>
- Sarkar S (2012) Evolution of relief in the Himalayan Foreland of West Bengal, India. *Geogr Thought* 10:1–11
- Starkel L, Sarkar S, Soja R, Prokop P (2008) Present day evolution of the Sikkimese–Bhutanese Himalayan Piedmont, Instytut Geografii Przestrzennego, pp 23–146
- Syvitski JPM, Vorosmarty CJ, Kettner AJ (2005) Impact of humans on the flux of terrestrial sediment to the global coastal ocean. *Science* 308(5720):376–380
- Tal M, Gran K, Murray AB, Paola C, Hicks DM (2004) Riparian vegetation as a primary control on channel characteristics in multi-thread rivers. *Water Sci Appl (Am Geophys Union)* 8:43–57. <https://doi.org/10.1029/008WSA04>
- Tamang L (2013) Effects of boulder lifting on the fluvial characteristics of lower Balason basin in Darjeeling district west Bengal. Unpublished Ph.D. thesis, Department of Geography and Applied Geography, NBU, under the supervision of Deepak Kumar Mandal, pp 66–87
- Terry JP, Goff J (2014) Megaclasts: Proposed revised nomenclature at the coarse end of the Udden–Wentworth grain-size scale for sedimentary particles revised nomenclature at the coarse end of the Udden–Wentworth Scale. *J Sediment Res* 84(3):192–197. <https://doi.org/10.2110/jsr.2014.19>
- Tonnen WHJ, Winkels TG, Cohen KM, Prins MA, Middelkoop H (2014) Lower Rhine historical flood magnitudes of the last 450 years reproduced from grain size measurements of flood deposits using End Member Modeling. *CATENA* 130:68–75. <https://doi.org/10.1016/j.catena.2014.12.004>
- Varga G, Újvári G, Kovács J (2019) Interpretation of sedimentary (sub) populations extracted from grain size distributions of Central European loess–paleosol series. *Quat Int* 502:60–70. <https://doi.org/10.1016/j.quaint.2017.09.021>
- Walsh JP, Nittrouer CA (2009) Understanding fine grained river sediment dispersal on continental margins. *Mar Geol* 263(1–4):34–45. <https://doi.org/10.1016/j.margeo.2009.03.016>
- Wentworth CK (1922) A scale of grade and class terms for Clastic sediments. *J Geol* 30:377–387
- Wheatcroft RA, Sommerfield CK, Drake DE, Borgeld JC, Nittrouer CA (1997) Rapid and widespread dispersal of flood sediment on the northern California margin. *Geology* 25(2):163–166. <https://doi.org/10.1130/0091-7613>
- Wolman MG (1954) A method of sampling coarse river-bed material. *Trans Am Geophys Union* 35(6):951–956
- Wu L, Krijgsman W, Liu J, Li C, Wang R, Xiao W (2020) CFLab: A MATLAB GUI program for decomposing sediment grain size distribution using Weibull functions. *Sediment Geol* 398:105590. <https://doi.org/10.1016/j.sedgeo.2020.105590>
- Yang H, Shi C (2019) Sediment grain-size characteristics and its sources of ten wind-water coupled erosion tributaries (the Ten Kongduis) in the Upper Yellow River. *Water* 11(1):2–15. <https://doi.org/10.3390/w11010115>
- Yim O, Ramdeen KT (2015) Hierarchical cluster analysis: comparison of three linkage measures and application to psychological data. *Quant Methods Psychol* 11(1):8–19. <https://doi.org/10.20982/tqmp.11.1>
- Yu SY, Colman SM, Li L (2016) BEMMA: a hierarchical Bayesian end-member modeling analysis of sediment grain-size distributions. *Math Geosci* 48(6):723–741. <https://doi.org/10.1007/s11004-015-9611-0>
- Zhang J, Ding Z, You J (2014a) The joint probability distribution of runoff and sediment and its change characteristics with multi-time scales. *J Hydrol Hydromech* 62(3):218–225. <https://doi.org/10.2478/johh-2014-0024>
- Zhang K, Chanpura RA, Mondal S, Wu CH, Sharma MM, Ayoub JA, Parlar M (2014b) Particle size distribution measurement techniques and their relevance or irrelevance to sand control design. In: SPE international symposium and exhibition on formation damage control. Society of Petroleum Engineers. <https://doi.org/10.2118/168152-MS>
- Zhou X, Li A, Jiang F, Lu J (2015) Effects of grain size distribution on mineralogical and chemical compositions: a case study from size-fractional sediments of the Huanghe (Yellow River) and Changjiang (Yangtze River). *Geol J* 50(4):414–433. <https://doi.org/10.1002/gj.2546>
- Zhang J, Zhang LY, Du M, Zhang W, Huang X, Zhang YQ, Li YW (2016) Identifying the major air pollutants base on factor and cluster analysis, a case study in 74 Chinese cities. *Atmos Environ* 144:37–46. <https://doi.org/10.1016/j.atmosenv.2016.08.066>
- Zhang Q, Zhu C, Yang LT, Chen Z, Zhao L, Li P (2017) An incremental CFS algorithm for clustering large data in industrial internet of things. *IEEE Trans Ind Inf* 13(3):1193–1201. <https://doi.org/10.1109/TII.2017.2684807>
- Zhang X, Zhou A, Wang X, Song M, Zhao Y, Xie H, Chen F (2018) Unmixing grain-size distributions in lake sediments: a new method of endmember modeling using hierarchical clustering. *Quat Res* 89(1):365–373. <https://doi.org/10.1017/qua.2017.78>
- Zhang T, Wang J, Cui C, Li Y, He W, Lu Y, Qiao Q (2019) Integrating geovisual analytics with machine learning for human mobility pattern discovery. *ISPRS Int J Geo-Inf* 8(10):434. <https://doi.org/10.3390/ijgi8100434>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”). Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.

These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.

We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.

While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval , sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.

These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.

To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.

Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.

If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com