Understanding meteorological and physio-geographical controls of variability of flood event classes in headstream catchments of China

Yongyong Zhang[1], Yongqiang Zhang[1], Xiaoyan Zhai[2], Jun Xia[3,1], Qiuhong Tang[1], Wei Wang[1], Jian Wu[1], Xiaoyu Niu[1], Bing Han[1]

[1]Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, 100101, China.
[2]China Institute of Water Resources and Hydropower Research, Beijing, 100038, China.
[3]State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan, 430072, China.

Correspondence to: Yongyong Zhang (zhangyy003@igsnrr.ac.cn)

Abstract. Classification is beneficial for understanding flood variabilities and their formation mechanisms from massive flood event samples for both flood scientific research and management purposes. Our study investigates comprehensive manageable flood event classes from 1446 unregulated flood events in 68 headstream catchments of China using the hierarchical and partitional clustering methods. Control mechanisms of meteorological and physio-geographical factors (e.g., meteorology, land cover and catchment attributes) on spatial and temporal variabilities of individual flood event classes are explored using constrained rank analysis and Monte Carlo permutation test. Results show that we identify five robust flood event classes, i.e., moderately, highly, and slightly fast floods, as well as moderately and highly slow floods, which accounts for 24.0%, 21.2%, 25.9%, 13.5% and 15.4% of total events, respectively. All the classes are evenly distributed in the whole period, but the spatial distributions are quite distinct. The fast flood classes are mainly in the southern China, and the slow flood classes are mainly in the northern China and the transition region between southern and northern China. The meteorological category plays a dominant role in flood event variabilities, followed by catchment attributes and land covers. Precipitation factors, such as volume and intensity, and aridity index are the significant control factors. Our study provides insights into flood event variabilities and aids in flood prediction and control.

1 Introduction

Flood events usually show tremendous spatial and temporal variabilities in behavior due to heterogeneities in meteorological and underlying surface conditions over large basins or entire regions (e.g., county, continent and world) (Berger and Entekhabi, 2001). Existing studies provide insights on impacts of changes in meteorological or underlying surface conditions on specific flood metrics (e.g., magnitude, peak, timing or seasonality) and their changes using trend separation method, correlation testing, mathematical modelling, and so on (Berghuijs *et al*., 2016; Tarasova *et al*., 2018; Liu *et al*., 2020; Wang *et al*., 2024). However, all of these studies are implemented at event scale or in catchments with certain landscapes and climates, which are insufficient for the comprehensive flood change investigation and generalized results (Tarasova *et al*., 2019; Zhang *et al*., 2020). Flood event similarity analysis is beneficial to investigate comprehensive dynamic characteristics of flood events in space and time by grouping massive heterogenous events into some manageable classes with significantly statistical differences of flood responses (e.g., great or small floods, fast or slow floods, rain or snowmelt floods) (Brunner, 2018). Flood event class determines hydrological response characteristics, longitudinal and lateral transfers of energy and material, and structures and functions of riverine ecosystems (Arthington *et al*., 2006; Poff and Zimmerman, 2010). The class also directly determines flood disaster losses for human society and affects the strategy formulations of flood control and management (Hirabayashi *et al*., 2013; Jongman *et al*., 2015). Hence, for both flood scientific research and management purposes, it is fundamentally important to identify the flood event classes and their formation mechanisms (Sikorska *et al*., 2015).

Inductive and deductive approaches are reported for the flood event similarity classification according to the clustering objectives (Olden *et al*., 2012). The inductive approach directly focuses on the shape similarity of flood events by clustering the response characteristics extracted from the flood event hydrographs. The response characteristics include magnitude, frequency, duration, timing and seasonality, variability metrics, which are considered as the critical components to characterize the entire range of flood events (Poff *et al*., 1997; Kuentz *et al*., 2017;Zhang et al., 2020). The reported flood event classes are the fast events with steep rising and falling limbs, the slow events with both elongated rising and falling limbs, the sharp or fast flood event, the flash flood (Kuentz *et al*., 2017; Brunner *et al*., 2018; Zhai *et al*., 2021; Zhang *et al*., 2020). The deductive approach mainly focuses on the similarity of environmental factors assumed to control flood events, such as meteorological variables (e.g., storm intensity, duration and snowmelt) and physio-geographical conditions (e.g., soil moisture, land cover and topography) (Merz and Blöschl, 2003; Ali *et al*., 2012; Brunner *et al*., 2018; Zhang *et al*., 2022). The reported flood event classes are the long-rain floods, short-rain floods, flash floods, rain-on-snow floods, and snowmelt floods (Merz and Blöschl, 2003; Sikorska *et al*., 2015; Brunner *et al*., 2018; Zhang *et al*., 2022). However, the control relationships of environmental factors on flood event shapes are not well defined so that the identified classes are not exactly helpful to investigate the flood change patterns at event scale. Therefore, it is a challenge to better understand the formation mechanisms of individual flood event classes.

The main procedure of existing flood event classification was to cluster the similarity of flood event attributes (e.g., flood response characteristics or control factors) across the spatial and temporal scales. According to the classification procedure, there are two widely-adopted approaches, namely the tree clustering methods (e.g., decision tree, regression tree, fuzzy tree and random forest) (Sikorska *et al*., 2015; Brunner *et al*., 2017) and the non-tree clustering methods (e.g., single linkage, complete linkage, average linkage, centroid linkage, ward linkage, k-mean, k-medoids) (Zhang *et al*., 2020; Zhai *et al*., 2021). The tree clustering methods as the hard clustering methods, are implemented to binarily split all the flood events successively into smaller classes of similar flood events according to the thresholds of flood response metrics until obtaining final classes (Sikorska *et al*., 2015; Brunner *et al*., 2017). The classification results could be applicable to other basins and the flood response characteristics of different studies would be directly comparable if the same thresholds are adopted. However, these methods assume that the boundaries of flood response metrics in different classes are clear and the thresholds of flood response metrics should be predefined and should not overlap among different classes (Olden *et al*., 2012; Sikorska *et al*., 2015; Zhai *et al*., 2021). Additionally, the classification is very sensitive to the thresholds, whose small changes would cause different flood event classes (Olden *et al*., 2012; Sikorska *et al*., 2015). Therefore, it will be difficult to define the thresholds clearly to get robust classification performance. The non-tree clustering methods as the soft clustering methods, are implemented to directly split all the flood events according to different division rules of the comprehensive similarity measures of flood event shapes or metrics (Olden *et al*., 2012; Zhang *et al*., 2020). The class boundaries of flood response metrics are not clear, which are mainly based on sufficient of heterogeneous flood events (Sikorska *et al*., 2015). The flood response characteristics of individual classes were usually qualitatively described to distinguish the differences among classes (Olden *et al*., 2012; Tarasova *et al*., 2019; Zhang *et al*., 2020). Therefore, the classification results obtained from different flood event samples are still difficult to quantitatively compare even though the flood response characteristics or hydrographs in the certain class are similar (e.g., high or low, fast or slow floods) (Zhang *et al*., 2024). However, these methods were widely-used due to their ease of use (Olden *et al*., 2012; Tarasova *et al*., 2019; Zhang *et al*., 2020). Additionally, the determinations of clustering method and final cluster number are subjective in most existing studies, and the assessment of clustering performance is usually unavailable (Olden *et al*., 2012; Sikorska *et al*., 2015; Brunner *et al*., 2017). Therefore, robustness of flood event classification should be further explored.

The main aim of this study is to investigate the flood event similarity and the control mechanisms of meteorological and physio-geographical factors in space and time at class scale across China. Over one thousand unregulated flood events at 68 heterogeneous catchments with wider meteorological and physio-geographical conditions are selected for our study. The specific objectives are as follows:

(i) to determine the optimal flood event classes by comparing multiple classification performance criteria of both the hierarchical and partitional clustering methods;

(ii) to identify the main flood response characteristics of individual classes and their spatial and temporal variabilities;

(iii) to quantify the effects of meteorological and physio-geographical factors on the variabilities of individual flood event classes.

This study provides more comprehensive insights into meteorological and physio-geographical controls of variabilities of flood event classes at large scale, and provides the mechanism supports for predicting flood event classes.

2 Study area and data sources

According to the Köppen-Geiger climate classification (Peel *et al*., 2007), China has diverse climate types, including alpine tundra climate (ET for Köppen-Geiger codes), tropical climate (A), arid, steppe and cold climate (BSk), arid, desert and cold

(BWk), cold without dry season (Df), cold with dry winter (Dw), temperate without dry season (Cf) and temperate with dry winter (Cw). Most Köppen-Geiger climate types in China (i.e., A, Dw, Cf and Cw) are controlled by the southeast and southwest monsoons in the summer with temperate and humid climates and the northwestern and northeastern monsoons in the winter with cold and dry climates. In these monsoon controlled climate types, the mean annual precipitation was 365-2654 mm with a mean of 1184 mm, of which over 65% fell between May and September according to the gauged daily precipitation observations from 2001 to 2020 in these regions. This led to frequent flooding and thus the region in the monsoon controlled climate types is usually considered as the flood-prone area of China (China Institute of Water Resources and Hydropower Research and Research Center on Flood and Drought Disaster Prevention and Reduction, the Ministry of Water Resources, 2021). In the last decade, flooding occurred in 455 rivers annually, which affected 822 million people and averaged over 10 billion US dollars (Ministry of Water Resources of the People's Republic of China, 2020a).

Sixty-eight headstream stations spread across the flood-prone areas were selected with catchment areas ranging from 21 km^2 to 4830 km^2, which were in all the monsoon controlled climate types of China, except tropical climate in the islands (i.e., A) (Figure 1). Most catchments had large forest coverage, with mean area percentages of 67.0%, particularly in the Yangtze (69.9%) and Pearl (68.7%) River Basins. A total of 1446 unregulated flood events with hourly time steps were collected from the Hydrological Yearbooks of the Songliao, Yellow, Huaihe, Yangtze, Southeast and Pearl River Basins over the period 1993—2015. The event was extracted following the Standard of Ministry of Water Resources of the People's Republic of China, i.e., Code for hydrologic data processing (SL/T 247—2020) (Ministry of Water Resources of the People's Republic of China, 2020b). The extracted flood events at the individual stations usually had the maximum flood peak or flood volume, isolated flood peak, continuous flood peaks, or flood peak after prolonged drought during the high and normal flow years (Ministry of Water Resources of the People's Republic of China, 2020b). There were 53 events at four stations in the Songliao River Basin, 104 events at four stations in the Yellow River Basin, 215 events at 13 stations in the Huaihe River Basin, 844 events at 38 stations in the Yangtze River Basin, 90 events at five stations in the Southeast River Basin, and 140 events at four stations in the Pearl River Basin. No less than 10 flood events were collected for every station to ensure the representativeness. The densities of flood events and gauges in the Huaihe River Basin and Southern China were much greater than those in the Songliao and Yellow River Basins in the Northern China because of the higher occurrences of flood events (Table S1 in the Supplement) (China Institute of Water Resources and Hydropower Research and Research Center on Flood and Drought Disaster Prevention and Reduction, the Ministry of Water Resources, 2021).

Meteorological, catchment and land cover data sources were collected together to calculate the potential meteorological and physio-geographical control factors and assess their contributions on the spatial and temporal variabilities of flood event classes. The meteorological data sources were the synchronous hourly precipitation events which were also extracted from the Hydrological Yearbooks, and the daily precipitation, maximum and minimum temperature observations from 1993 to 2015 at meteorological stations within or around the catchments which were downloaded from the China Meteorological Data Sharing

Service System. All the meteorological stations in the buffer zone with a radius of 100 km of every catchment centers were selected. The station number was 466 in total and no less than eight stations for each catchment. The daily meteorological variables were interpolated to the catchment by the inverse distance weighting method. The geographic information system data were collected to extract the attributes of catchment and land cover. The detailed data were the digital elevation model with a spatial resolution of 30 m×30 m and land covers in six periods (i.e., 1990, 1995, 2000, 2005, 2010 and 2015) with a spatial resolution of 30 m×30 m, all of which were downloaded from the Data Center of Resources and Environmental Science, Chinese Academy of Sciences. All these control factors well represented the meteorological and underlying surface conditions of individual catchments because all these flood events were captured satisfactorily by the catchment hydrological model developed using these factors (Zhang *et al*., 2024).

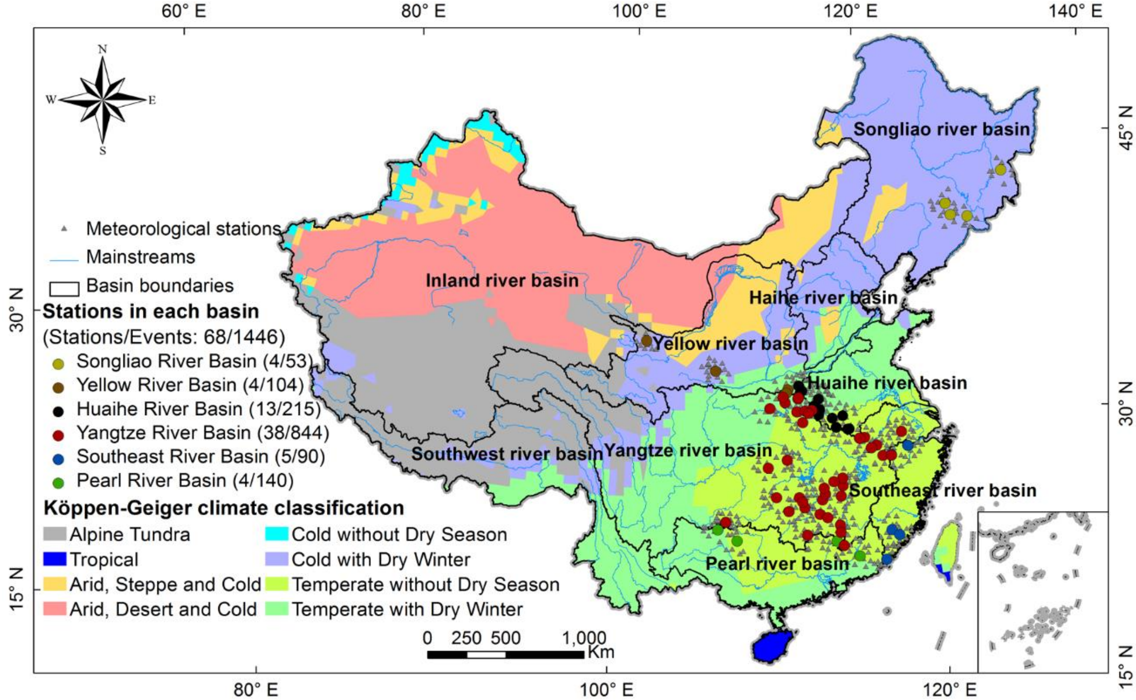


Figure 1. Spatial distributions of all the selected flood events and their corresponding climate types

3 Methods

3.1 Flood response metrics

The flood classification in our study mainly focuses on the detailed response characteristics of flood hydrographs by the inductive approach. The magnitude, variability, timing, duration, and rate of changes are widely-accepted as the main five components to characterize the entire flood events (Poff *et al*., 2007) and thus are also adopted to characterize the detailed flood responses in our study. Additionally, flood peak number is one of the most important metrics for flood control (Aristeidis *et al.*, 2010; Rustomji *et al.*, 2009). Therefore, nine metrics are used to fully characterize the response of flood events. There are the magnitude (total flood volume: R, maximum flood peak: Q_{pk}), variability (coefficient of variation: CV), timing (timings of flood event and maximum flood peak: T_{bgn} and T_{pk}), duration (flood event duration: T_{drn}), rate of changes (mean rates of positive and negative changes: RQ_r and RQ_d) and flood peak number (N_{pk}) (Arthington *et al.*, 2006; Kennard *et al.*, 2010; Poff *et al.*, 2007; Zhang *et al.*, 2012). T_{bgn} is characterized using the circular statistical approach which translates the calendar date into the polar coordinates on the circumference of a circle, and is beneficial to distinguish the seasonal pattern (Fisher, 1993; Dhakal *et al.*, 2015). Table 1 summarizes the definitions of all the selected flood components.

Table 1. Metrics used to characterize flood responses in our study

Components	Metrics	Abbreviations	Units	Equations	References
Magnitude	Total flood volume	R	mm·day^{-1}	$R = 86.4 \cdot \sum_{t=TF_{bgn}}^{TF_{end}} Q_t / A$	Fisher, 1993; Black and Werritty, 1997; Poff *et al.*, 2007; Villarini, 2016; Hall and Blöschl, 2018; Zhang *et al.*, 2020
	Maximum flood peak	Q_{pk}	mm·day^{-1}	$Q_{pk} = \max(86.4 \cdot Q_t) / A$	
Variability	Coefficient of variation	CV	-	$CV = \sigma / Q_{av}$	
Timing	Ratio of beginning date of flood event in the calendar year using circular statistics	T_{bgn}	radian	$T_{bgn} = 2\pi \cdot TF_{bgn} / TD$	
	Ratio of occurrence time of maximum flood peak to flood duration	T_{pk}	%	$T_{pk} = TF_{pk} / T_{drn} \cdot 100$	
Duration	Duration of flood event	T_{drn}	h	$T_{drn} = 24 \cdot (TF_{end} - TF_{bgn} + 1)$	
Rate of changes	Mean rate of positive changes	RQ_r	h^{-1}	$RQ_r = \frac{(Q_{pk} - Q_{bgn}) / Q_{av}}{(TF_{pk} - TF_{bgn} + 1) \cdot 24}$	
	Mean rate of negative changes	RQ_d	h^{-1}	$RQ_d = \frac{(Q_{pk} - Q_{end}) / Q_{av}}{(TF_{end} - TF_{pk} + 1) \cdot 24}$	
Number	Number of peaks during the event	N_{pk}	-		Aristeidis *et al.*, 2010; Zhai *et al.*, 2021

Note: Q_t is the flood magnitude on day t (m^3·s^{-1}); Q_{av} is the mean flood magnitude (m^3·s^{-1}); Q_{bgn} and Q_{end} are flood magnitudes at the beginning and end of event (m^3·s^{-1}), respectively; σ is the standard deviation of flood magnitude (m^3·s^{-1}); TD is the total days of the calendar year (day), i.e., 365 for common year or 366 for leap year; TF_{bgn} and TF_{end} are the beginning and end dates of flood events; TF_{pk} is the occurrence date of maximum flood peak; A is the catchment area (km^2) ; 86.4 is the unit conversion factor from m^3·s^{-1}·km^{-2} to mm.

3.2 Flood event classification

High dimensionality and multicollinearity exist among flood response metrics and affect the flood event classification when a large number of metrics are considered (Olden *et al*., 2012; Zhang *et al*., 2012). Dimensionality reduction is to transform the high dimensional metrics into a few independent composite metrics without losing the metric information, and to reveal the major similarity characteristics among the metrics. Here, principal component analysis is used to obtain a few principal components (*PCA*) based on the orthogonal transform. If the cumulative variance is over 85% of the total explained variances of all the flood response metrics, the first *m PCA*s are selected for classification. The main flood response metrics in the individual *PCA*s were determined according to the load coefficient matrix. If the load coefficient is over 0.45, the corresponding flood response metric are considered to be highly correlated with the *PCA*.

Subsequently, both the hierarchical (Ward's) and partitional (*k*-medoids) clustering methods are used to cluster flood events based on the similarity of the selected *PCA*s. Euclidean distance is the distance measure used. Twenty-two criteria are used to assess the classification performance and determine the best number of clusters, i.e., KL, CH, Hartigan, CCC, Scott, Marriot, TrCovW, TraceW, Friedman, Silhouette, Ratkowsky, Ball, Ptbiserial, Dunn, Rubin, Cindex, DB, Duda, Pseudot2, McClain, SDindex and SDbw (Table S2 in the Supplement) (Charrad et al., 2014). The greater values of the first fourteen indexes (i.e., KL to Dunn) or the smaller values of the rest eight criteria (i.e., Rubin to SDbw) indicate the better classification. If the best criteria number is the largest in a certain cluster number, the cluster number is optimal and the corresponding clustering method is also selected.

All the multivariable statistical analyses are implemented using R software (version 3.1.1) (R Development Core Team, 2010), involving the `aov`, `cor` and `princomp` functions in stats Package (version 4.1.3) for independence test, linear correlation test and principal component analysis, respectively (Mardia *et al*., 1979), the `hcluster` function in amap Package (version 0.8-18) for hierarchical cluster analysis (Antoine and Sylvain, 2006), the `clara` function in cluster Package (version 2.1.3) for *k*-medoids cluster analysis (Kaufman and Rousseeuw, 1990), the `NbClust` function in NbClust Package (version 3.0.1) for the optimal class number determination and classification performance assessment (Charrad *et al*., 2014).

3.3 Control mechanisms of meteorological and physio-geographical factors on the variabilities of flood event classes

3.3.1 Meteorological and physio-geographical factors

The meteorological (e.g., storm intensity, timing and duration, evapotranspiration volume) and physio-geographical factors (e.g., land covers and catchment attributes) directly affect the flood generation and routing processes, which thus cause the diversity of flood event shapes (Ali *et al*., 2012; Brunner *et al*., 2018; Merz and Blöschl, 2003; Zhang *et al*., 2022). The potential control factors are selected as many as possible to investigate the control mechanisms on the variability of flood event classes according to the existing studies and the total number is 34 meteorological, catchment and land cover factors in all the catchments (Table 2). In the meteorological category, 17 factors related to precipitation, potential evapotranspiration and aridity index are selected, including the amounts, intensities and timing factors during flood events, in the antecedent period and at annual scale. All the precipitation factors during the flood events are extracted using the hourly precipitation observations. The precipitation factors at daily or annual scale are extracted using the daily precipitation observations. The potential evapotranspiration at daily or annual scale is estimated using the Hargreaves method (Hargreaves and Samani, 1982), and the aridity index is the ratio of potential evapotranspiration to precipitation. All these factors mainly affect the flood yield processes (Merz and Blöschl, 2003; Aristeidis *et al*., 2010; Zhang *et al*., 2022).

For the physio-geographical factors, the 10 catchment attributes are selected, including catchment location, area, elevation and slope, river density and slope. All these factors mainly affect the flood yield and routing processes (Ali *et al*., 2012; Kuentz *et al*., 2017). Seven land cover factors are selected, including the area fractions of paddy, dryland, forest, grassland, water, urban and rural area to the total catchment, respectively for the seven land cover periods. All these factors mainly affect the flood yield and overland routing processes (Kuentz *et al*., 2017; Zhai *et al*., 2021). All these factors are extracted using the `Hydrology` and `Zonal` functions of the Spatial Analyst Tools in the ArcGIS Desktop (version 10.0).

Table 2. Meteorological and physio-geographical factors in our study

Factor categories		Factors	Data sources	Flood event effects
Meteorology	Precipitation	• pcp_ant: cumulative amount in the antecedent seven days (mm); • pcp_dur:total amount during the flood event (mm); • pcp_av: mean amount during the flood event (mm hr^{-1}); • pcp_max: maximum intensity during the flood event (mm hr^{-1}); • pcp_max: maximum intensity during the flood event (mm hr^{-1}); • pcp_Tbeg: precipitation timing; • pcp_Tdur: precipitation duration (days); • pcp_ann: annual mean amount (mm); • pcp_year: amount in the year when the flood event happens (mm)	Hourly precipitation in hydrological yearbooks; daily precipitation at 466 meteorological stations	Flood yield process
	Potential evapotranspiration	• pet_ant: cumulative amount in the antecedent seven days (mm); • pet_dur: total amount during the flood event (mm) • pet_max: maximum intensity during the flood event (mm hr^{-1}) • pet_ann: annual mean amount (mm); • pet_year: amount in the year when the flood event happens (mm)	Daily maximum and minimum temperature at 466 meteorological stations	Flood yield process

	Aridity index	• SPEI_ant: mean value in the antecedent seven days; • SPEI_dur: mean value during the flood event ; • SPEI_ann: annual mean value; • SPEI_year: mean value in the year when the flood event happens	Daily maximum and minimum temperature at 466 meteorological stations	Flood yield process
Physio-geography	Locations	• Longitude: longitude of catchment center • Latitude: latitude of catchment center	Global positioning system	Meteorological conditions
	Catchment attributes	• Slope: catchment slope (%); • Area: catchment area (km^2); • Length: catchment slope length (km); • Elevation: average elevation of catchment (m); • MaxiElev: maximum elevation of catchment (m);	Digital elevation model (size: 30 m×30 m)	Flood yield and overland routing processes
	River attributes	• Rivden: river density (km/km^2); • RivSlope: river slope (%); • Rwd: ratio of river width to depth (m/m);	Digital elevation model (size: 30 m×30 m)	Flood routing processes in river system
	Land covers	• Rpaddy: area fraction of paddy to catchment (%); • Rdryland: area fraction of dryland to catchment (%); • Rforest: area fraction of forest to catchment (%); • Rgrass: area fraction of grass to catchment (%); • Rwater: area fraction of water to catchment (%); • Rurban: area fraction of urban to catchment (%); • Rrural: area fraction of unused land to catchment (%)	Land covers in 1990, 1995, 2000, 2005, 2010 and 2015 (size: 30 m×30 m)	Flood yield and overland routing processes

3.3.2 Effect quantifications of meteorological and physio-geographical factors

The constrained rank analysis is adopted to quantify the direct or combined effects of control factor categories on spatial and temporal variabilities of individual flood event classes for both the distributed and lumped analyses. The widely adopted methods of constrained rank analysis are the Redundancy Analysis (RDA) and the Canonical Correlation Analysis (CCA). The RAD is a linear model and the CCA is a unimodal model, both of which are the extended methods of principal component analysis combined with regression analysis. These methods have strong advantages to solve multiple linear regressions and interactions between dependent and independent variable matrixes which are transformed into a few independent composite factors (ter Braak, 1986; Legendre and Anderson, 1999), and are beneficial to quantify the effects of independent variable matrix on dependent variable matrix and to find the most important factors. Both methods have been commonly used in testing the multispecies response to environmental variables in the biological or ecological sciences (Legendre and Anderson, 1999), effects of physio-geographical factors and human activities on diffuse nutrient losses or water quality (Zhang *et al.*, 2016; Shi *et al.*, 2017), and so on. The constrained proportion is the percentage of explained variance by independent variable matrix to the total variance of dependent variable matrix, which is usually considered as the effect contribution of individual meteorological and physio-geographical factors or categories on total variabilities of flood event classes. If the contribution sum of individual factor effects is less than the entire contribution of all the factors, the interactive effects are among the factors and the difference between the summed and entire contributions is the combined contribution (Zhang *et al.*, 2016). The selection is based on the first axis length. The CCA is proposed when the first axis length is greater than 4.0, while the RDA is proposed when the first axis length is less than 3.0. Otherwise, both CCA and RDA are proposed (ter Braak, 1986; Zhang *et al*., 2020).

Furthermore, the Monte Carlo permutation test is adopted to test the statistical significance of control factors on the variability of flood event classes, and obtain the correlation coefficients (r) between flood response matrix and control factor matrix in the individual catchments (i.e., distributed analysis) and the entire region (i.e., lumped analysis), respectively. Because all the catchment factors do not change in the entire period, only the meteorological and land cover categories are considered for the correlation relationship test for the lumped analysis. The significant statistical interval is set as 95%, i.e., p=0.05. The above-mentioned analyses are implemented using the `envfit`, `decorana`, `rda`, `cca`, `permutest` functions in the vegan Package (version 2.5-7) of R software (version 3.1.1) (ter Braak, 1986; R Development Core Team, 2010).

4 Results

4.1 Flood event classification

By the tests of independence and linear correlation for all the flood response metrics, T_{bgn} is independent from R, RQ_r, RQ_d and N_{pk}; Q_{pk} is independent from T_{pk}; and N_{pk} is independent from RQ_r and RQ_d. Expect these independent metrics, all the other metrics have linear correlations with each other (Table S3 in the Supplement). By the principal component analysis, five independent PCAs are found with the total cumulative variance of 85.7%, which are greater than the threshold (80.0%) (Table 3). Thus, the first five PCAs are selected in our study. According to the load coefficient matrix, the first PCA is related with magnitude (R and Q_{pk}), variability (CV) and rates of changes (RQ_r and RQ_d) with the load coefficients of 0.61, 0.97, 0.47, 0.84 and 0.84, respectively, and all of these metrics explain 33.3% of total variances of flood response metrics. The second PCA is related with R, CV, T_{pk} and N_{pk} with the load coefficients of 0.51, -0.47 and 0.56, respectively, and all of these metrics explain 17.0% of total variances. The third PCA is mainly related with T_{drn} and T_{pk} with the load coefficients of -0.48 and 0.48, respectively, and all of these metrics explain 16.0% of total variances. The fourth and fifth PCAs are mainly related with timings (T_{bgn} and T_{pk}) of flood event and maximum flood peak with the load coefficients of 0.92 and 0.64, respectively. The explained variances are 10.8% and 8.6%, respectively.

Table 3. Loads coefficients of flood response metrics in the selected PCAs and their explained variances

Metrics	$PCA1$	$PCA2$	$PCA3$	$PCA4$	$PCA5$
R	0.61	0.51	0.30	-0.01	0.06
Q_{pk}	0.97	0.08	0.00	-0.02	0.02
CV	0.47	-0.47	0.42	-0.28	0.40
T_{bgn}	0.05	-0.22	0.26	0.92	0.16
T_{pk}	-0.07	0.56	-0.48	0.03	0.64

T_{drn}	-0.19	0.16	0.84	-0.13	0.20
RQ_r	0.84	-0.22	-0.06	-0.01	-0.18
RQ_d	0.84	0.05	-0.29	0.10	0.02
N_{pk}	0.07	0.77	0.30	0.10	-0.31
Variance (%)	33.3	17.0	16.0	10.8	8.6
Cumulative variance (%)			85.7		

Compared with the classification performance of these two clustering methods (i.e., the hierarchical and *k*-medoids methods) among individual optimal cluster numbers (Figure 2), the optimal criteria number is the largest when the cluster number is five (i.e., 22.7% of total) for the *k*-medoids clustering method. The optimal criteria are the CCC, TrCovW, Silhouette, Ratkowsky and PtBiserial with the values of -2.98, 1.39×10^{15}, 4.12×10^{6}, 0.20, 0.29 and 0.39, respectively. Therefore, the five clusters using the *k*-medoids clustering method are optimum for further analysis in our study. The flood event numbers in the individual classes are 347, 306, 195, 375 and 223, accounting for 24.0%, 21.2%, 13.5%, 25.9% and 15.4% of total events, respectively.

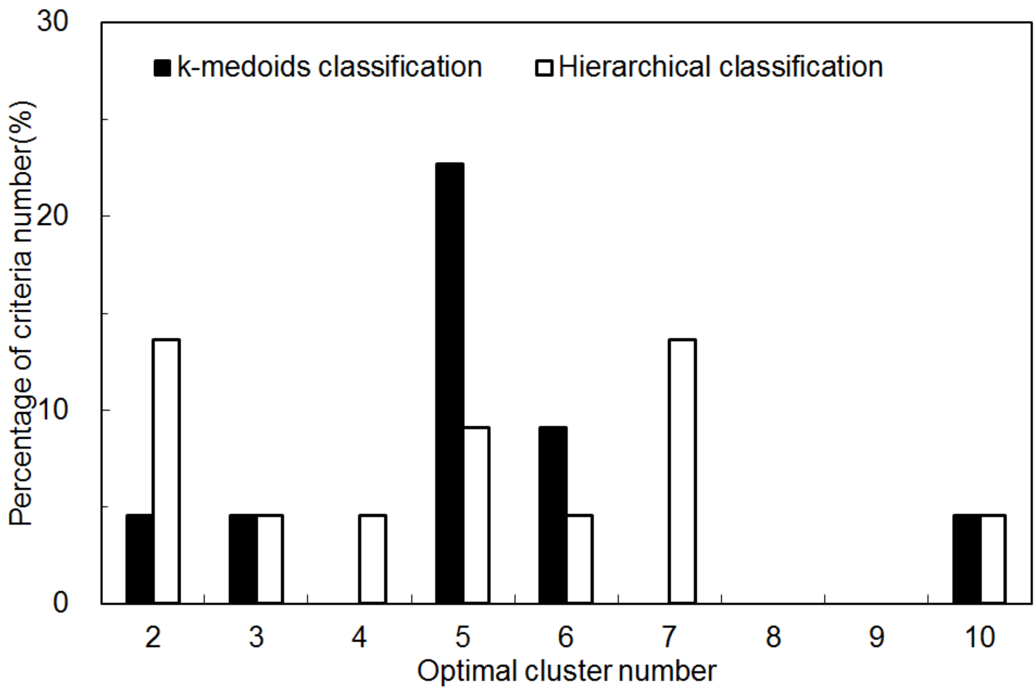


Figure 2. Classification performance comparisons between the hierarchical and *k*-medoids methods among individual optimal cluster numbers

4.2 Flood response characteristics of different classes

The value ranges of flood response metrics in different classes are presented in Figure 3 and Table S4 in the Supplement. For the magnitude metrics, both total flood volume (R) and maximum flood peak (Q_{pk}) variations are the same among different classes. The metric values in Class 3 are the largest, followed by Classes 5, 2, 1 and 4. For the variability metrics (CV), the events are the most variable in Class 5, and are slightly variable in the other Classes with the mean CV being less than 1.0, i.e., 0.90±0.26 (Class 1), 0.87±0.25 (Class 2), 0.86±0.26 (Class 4) and 0.84±0.22 (Class 3). For the timing and duration metrics (i.e., T_{bgn}, T_{drn} and T_{pk}), 73.2% of flood events in Class 1 occur before the wet season (i.e., January–May), and 58.5%, 67.7% and 57.0% of flood events in Classes 2, 3 and 5 occur in the earlier wet season (i.e., June–July), and 52.8% of flood events in Class 4 occur in the latter wet season (i.e., August–September). The mean duration (T_{drn}) is the longest in Class 5, followed by Classes 3 and 1. The mean T_{drn} values in Classes 4 and 2 are the shortest, i.e., 85.73±39.97 h and 83.82±41.20 h. The timings of maximum flood peaks (T_{pk}) are usually the largest in Class 2 with the mean of 50.6%±10.3%, which means that the flood peaks mainly occur in the middle or late stages of flood events. The flood peaks usually occur in the early stage of flood events in the other classes (i.e., Classes 1, 2, 4 and 5). Particularly in Class 3, the mean T_{pk} value is only 23.7%±13.6%.

For the rates of changes, RQ_r in most classes are much greater than RQ_d because the flood peaks usually occur in the early stage of flood events, except Class 2. The largest values of both RQ_r and RQ_d are in Class 3 because of the greatest flood peak. The smallest RQ_r values are mainly in Classes 2 because of the late occurrences of flood peaks, while the smallest RQ_d values are mainly in Class 5 because of the long durations of flood recession. For the flood peak number (N_{pk}), 71.2%, 69.9%, 76.5% and 77.1% of flood events has one flood peaks in Classes 1, 2, 4 and 5, respectively, and multiple flood peaks (i.e., two–four) exist in 94.4% of total flood events in Class 3, accounting for 33.8% (two peaks), 48.7% (three peaks) and 11.8% (four peaks), respectively.

Class 1 is for moderately fast flood events occurring before the wet season, characterized by a single peak and moderate duration, referred to as the "moderately fast flood event class" (Figure 4). Class 2 represents highly fast flood events with a single peak in the late stage and short duration, denoted as the "highly fast flood event class". Class 3 exhibits highly slow flood events during the latter part of the wet season, featuring multiple peaks and long duration, known as the "highly slow and multipeak flood event class". Class 4 reflects slightly fast flood events occurring in the latter wet season with a single peak and short duration, named the "slightly fast flood event class". Lastly, Class 5 displays moderately slow flood events with a single peak and long durations, designated as the "moderately slow flood event class".

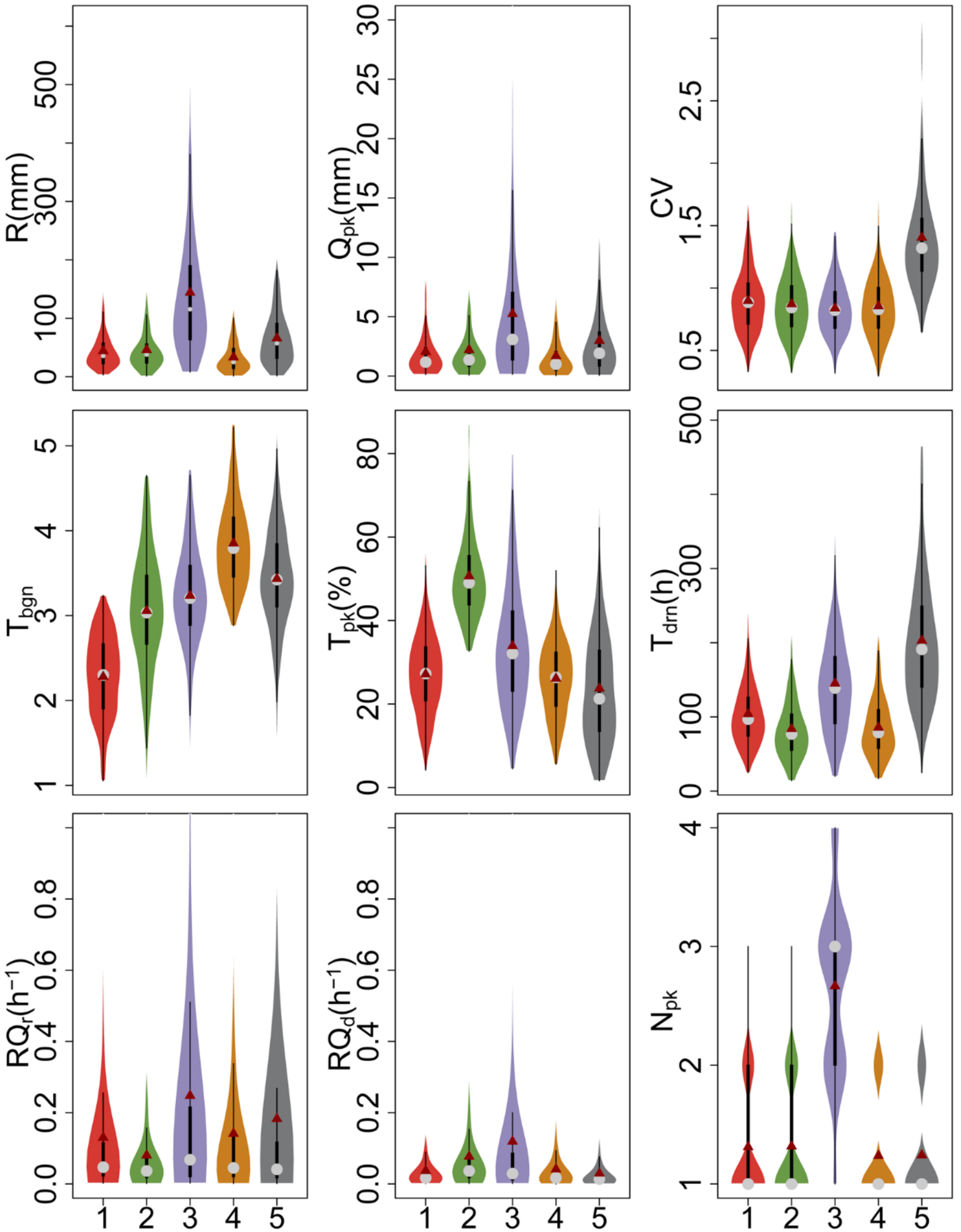
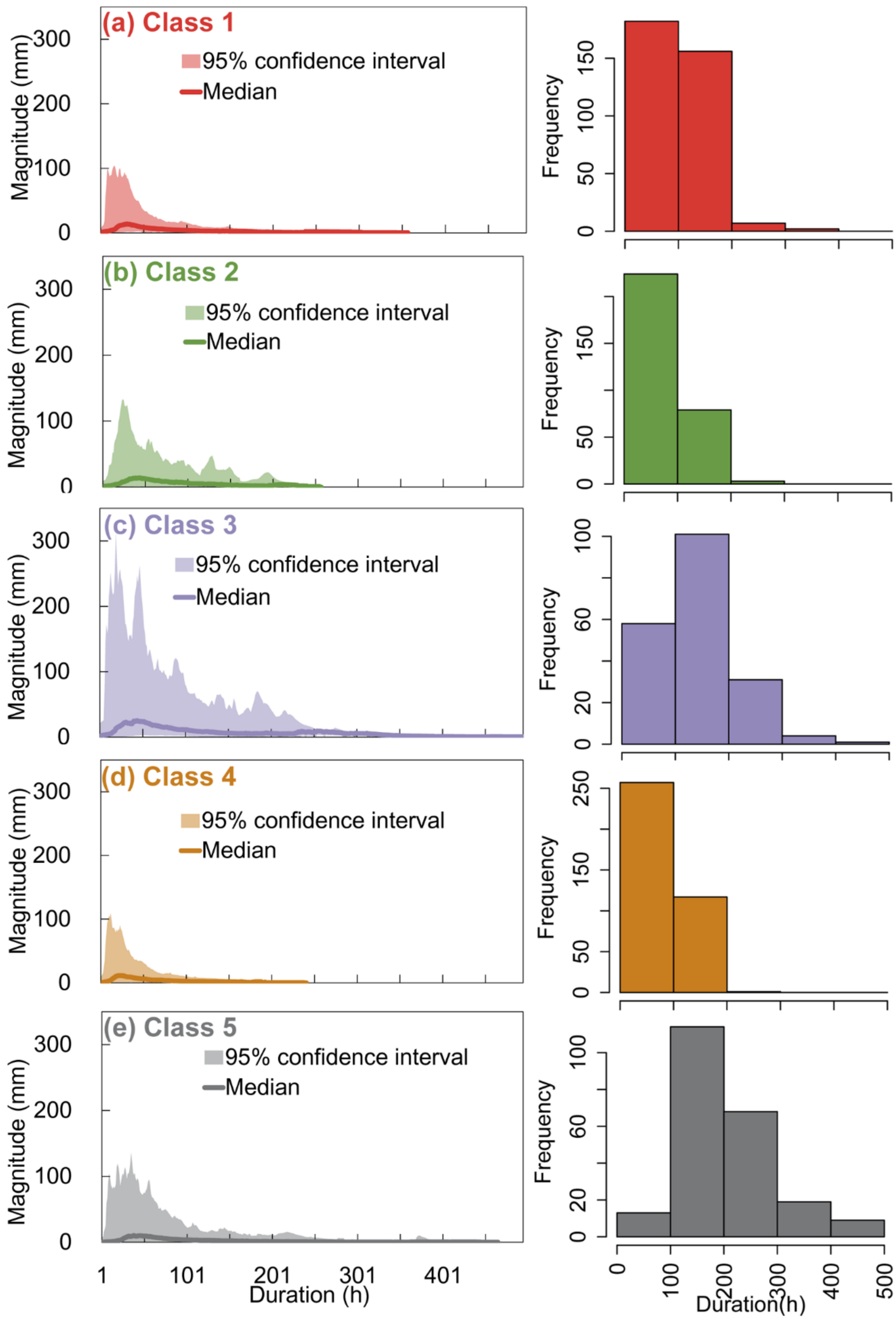


Figure 3. Variations of flood response metrics among Classes 1–5. The solid darkred dot and gray dot define the mean and 50th percentile values, respectively. Each black box means the 25th and 75th percentile values, and the vertical line defines the minimum and maximum values without outliers. The violin shape means the frequency distribution of flood response metric.



305 **Figure 4. Flood event distributions in the 95% confidence interval and their median, and their duration frequencies of Classes 1–5 (a–e)**

4.3 Spatial and temporal distributions of flood event classes

The spatial distributions of individual classes are showed in Figures 5 and S1, and Table S5 in the Supplement. The moderately fast flood event class (i.e., Class 1) is mainly in the Pearl and Yangtze River Basins, accounting for 37.1% (52/140) and 29.7% (251/844) of total events, respectively. Specifically, Class 1 is dominant in the Xiawan, Yanling and Songgao catchments in the Yangtze River Basin, and Hezikou catchment in the Pearl River Basin. The highly fast flood event class (i.e., Class 2) is mainly in the Pearl River Basin, accounting for 31.4% (44/140) of total events, particularly in the Xiaogulu catchment. The highly slow and multipeak flood event class (i.e., Class 3) is mainly in the Southeast River Basin, accounting for 42.2% (38/90) of total events, particularly in the Longshan catchment. The slightly fast flood event class (i.e., Class 4) is mainly in the Yellow and Songliao River Basins, accounting for 64.4% (67/104) and 60.4% (32/53) of total events, respectively. The most obvious catchments are Biyang in the Yangtze River Basin, Qiaotou and Luanchuan in the Yellow River Basin, Jingyu and Dongfeng in the Songliao River Basin. The moderately slow flood event class (i.e., Class 5) is mainly in the Huaihe River Basin, accounting for 47.4%(102/215) of total events, particularly in the Beimiaoji and Qilin catchments. Therefore, the Classes 1 to 3 are mainly in the Temperate without Dry Season climate region in southern China (Figure 1), the Class 4 is mainly in the Cold with Dry Winter climate region in northern China, and the Class 5 is mainly in the transition region between Temperate without Dry Season climate and Cold with Dry Winter climate.

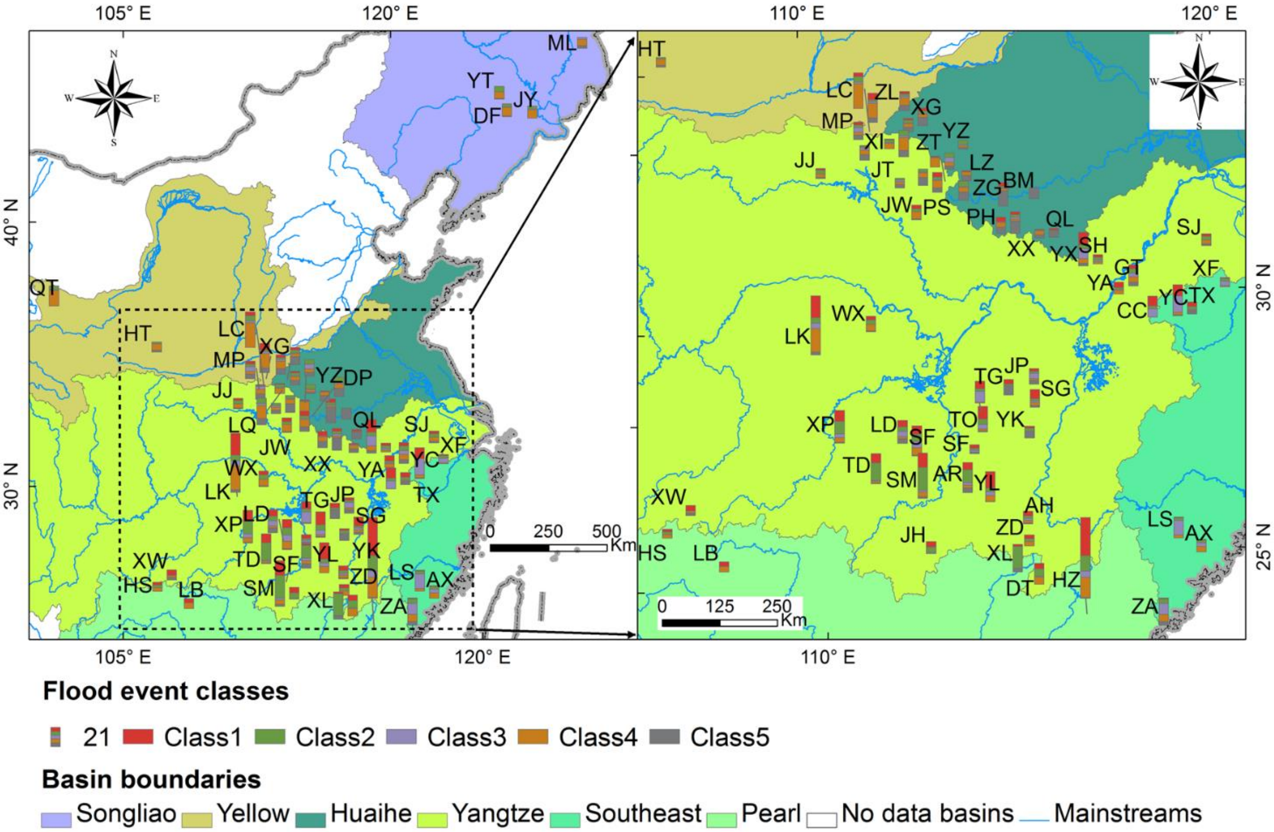


Figure 5. Spatial variabilities of individual flood event classes in major river basins

According to the interannual distributions of individual classes (Figure 6), all the classes are evenly distributed, whose annual mean percentages are 24.0±5.9%, 21.2±6.4%, 13.5±7.7%, 25.9±6.2%, and 15.4±12.5%, respectively. However, the interannual distributions of individual classes are quite distinct at different stations, particularly in the Songliao River Basin. In the headstream stations of Songliao River Basin, the dominant class is Class 4 with the annual mean percentage of 26.1±38.3% (*n*=32) though flood events are missed in several years due to the dry period. In the headstream stations of Yellow River Basin, the Class 4 is also dominant across the whole period with the annual mean percentage of 58.1±33.9% (*n*=67), particularly in 1994-1996, 1999 and 2007. In the headstream stations of Huaihe River Basin, the Class 5 gradually prevail with the annual mean percentage of 41.5±23.7% (*n*=102), particularly after 2007, whose percentage reaches 63.2±15.8% (*n*=79). The event numbers of both Classes 1 and 2 gradually decrease, accounting for 33.1±24.4% (*n*=11) and 8.7±7.1% (*n*=5) of annual flood events in the period of 1993-1999 and 2011-2015 for the Class 1, respectively, and 20.3±20.9% (*n*=9) and 2.7±1.3% (*n*=1) in the period of 1993-1999 and 2011-2015 for the Class 2, respectively. The explanations are that the total precipitation amount and duration probably increase due to the climate change (Dong *et al.*, 2011; Jin *et al.*, 2024). In the headstream stations of Yangtze River Basin, the Classes 1, 2 and 4 are dominant, accounting for 29.3±9.6% (*n*=251), 23.0±11.5% (*n*=197) and

21.1±7.0% (n=181) of annual mean flood events, respectively. Although the interannual changes of event numbers of Classes 1 (n=1–21), 2 (n=1–14) and 4 (n=1–16) are considerable, those of class percentages are relatively uniform except 2015. In the headstream stations of Southeast River Basin, the Class 3 gradually prevail after 2000 with the annual mean percentage of 46.2±32.5% (n=39). In the headstream stations of Pearl River Basin, the Class 1 is dominant with the annual mean percentage of 36.0±24.0% (n=52), but gradually shifts to Class 2 which accounts for 30.0±25.2% of annual mean flood events (n=40), particularly after 2008.

345

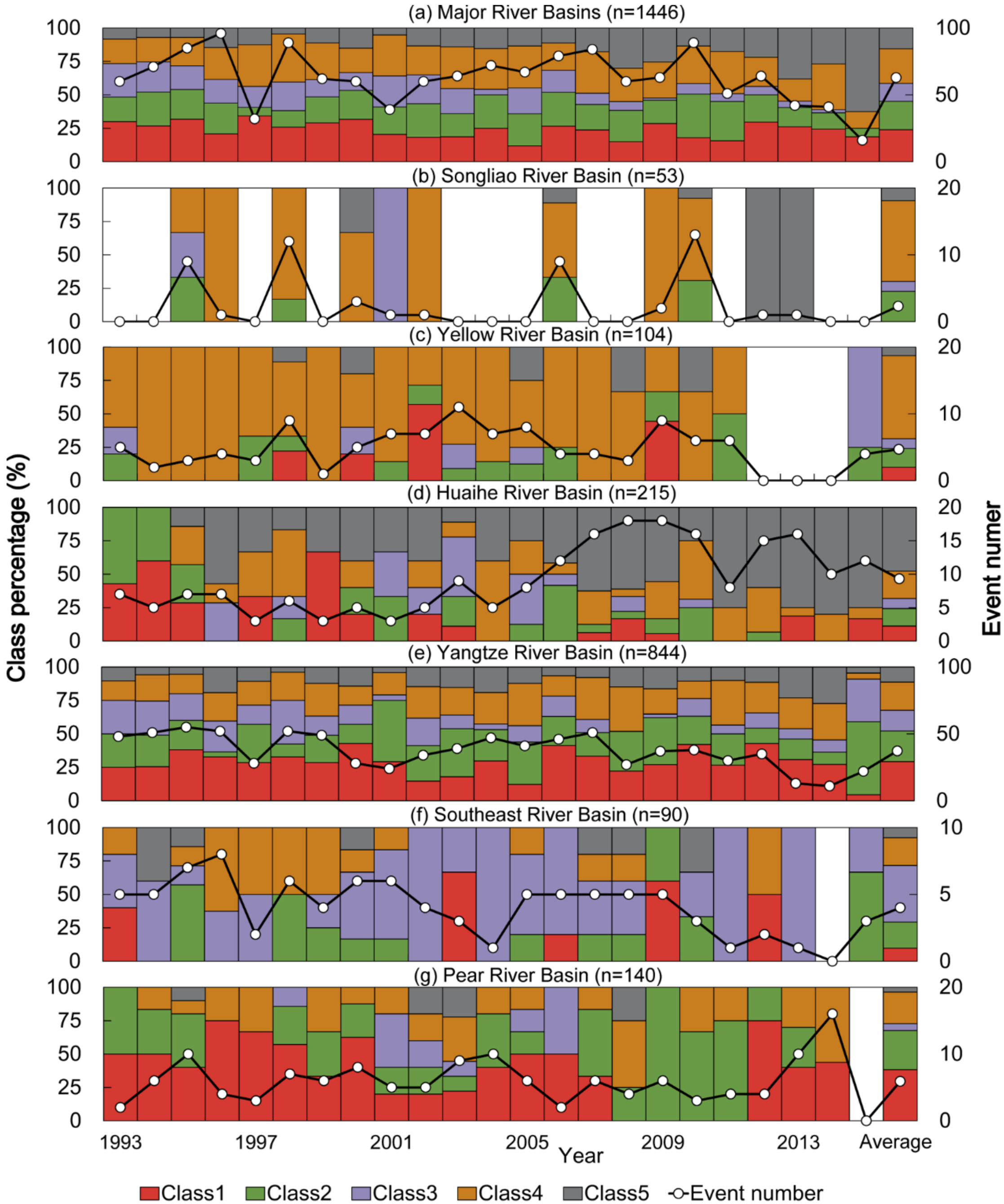


Figure 6. Interannual variabilities of individual flood event classes and their percentages in major river basins

4.4 Control mechanisms of meteorological and physio-geographical factors

4.4.1 Control factors and their contributions for the distributed analysis

According to the Monte Carlo permutation test between flood response matrix and control factor matrix (i.e., meteorological and land cover categories) in the individual catchments (Figures 7 and S2–5 in the Supplement), the factors only in the meteorological category are statistically significant for the temporal variabilities of flood events in all the classes, particularly the precipitation factors (e.g., amount, intensity) and aridity index during the events. Taking the Class 1 as an example, the total and mean precipitations, and aridity index during the event (r_{pcp_dur}=0.65–0.99, n=14; r_{pcp_av}=0.70–0.97, n=7; r_{SPEI_dur}=0.52–0.97, n=7) are the major control factors in 44.7% (17/38) of total catchments of the Yangtze River Basin, and Tunxi catchment of the Southeast River Basin and Hezikou catchment of the Pearl River Basin. The contributions of control factors are statistically significant only in the Liangshuikou and Hezikou catchments. In the Liangshuikou catchment, 96.3% of temporal differences are explained, in which the meteorological and land cover categories explain 92.5% and 3.8%, respectively. In the Hezikou catchment, 66.7% of temporal differences are explained, in which the meteorological category and the combined impact explain 49.4% and 17.3%, respectively.

In the Class 2, the significant control factors are in the catchments of Yangtze (18.4%, 7/38), Yellow (25%, 1/4) and Pearl (50%, 2/4) River Basins, particularly the total and mean precipitations, and aridity index during the event with the correlation coefficients of 0.61–0.99, 0.58–0.99 and 0.50–0.98, respectively. The contributions only in the Shimenkan, Tangdukou and Xiaogulu catchments are statistically significant with the total values of 90.7–96.8%. The contributions of meteorological category are the greatest with the values of 71.9–95.9%. In the Class 4, the significant control factors are in the catchments of Yellow (75%, 3/4), Songliao (50%, 2/4) and Pearl (50%, 2/4) River Basins, particularly the total precipitation during the event, and the aridity index in the corresponding year with the correlation coefficients of 0.53–1.00 and 0.45–0.93, respectively. The contributions only in the Liangshuikou and Hezikou catchments are statistically significant with the total values of 87.0–98.1%. The factors in the meteorological category also contribute the most considerably with the values of 76.8–82.1%. In the Classes 3 and 5, the contributions are not statistically significant in all the catchments because of the smaller numbers of flood events. However, several important control factors are also statistically significant in the catchments of Yangtze (26.3%, 10/38) and Southeast (40%, 2/5) River Basin for Class 3 (e.g., total and mean precipitations during the event with the correlation coefficients of 0.77–0.99 and 0.70–1.00, respectively), and Huaihe (61.5%, 8/13) and Yangtze (26.3%, 7/38) River Basin for Class 5 (e.g., the aridity index in the corresponding year and during the event, and the annual mean precipitation amount with the correlation coefficients of 0.62–0.86, 0.68–1.00 and 0.65–0.92, respectively).

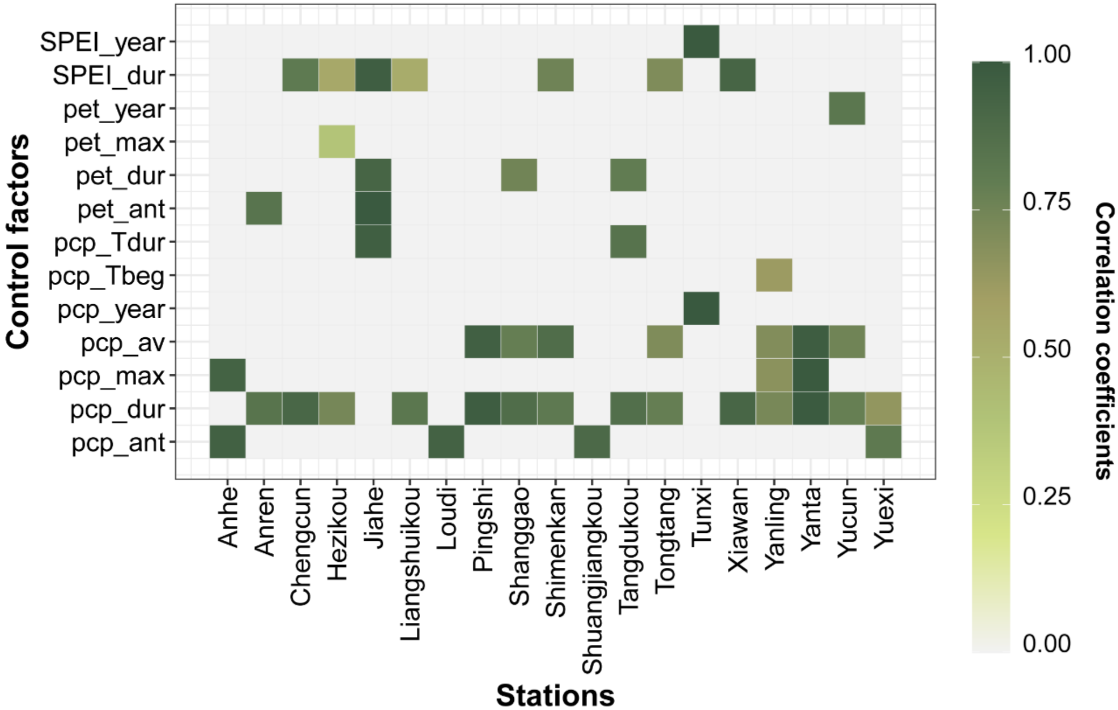


Figure 7. Significant control factors and their correlation coefficients for the temporal variabilities of flood event Class 1 in the individual catchments. The gray color means the control factor without statistical significance.

Note: Anhe, Anren, Chengcun, Jiahe, Liangshuikou, Loudi, Pingshi, Shanggao, Shimenkan, Shuangjiangkou, Tangdukou, Tongtang, Xiawan, Yanling, Yanta, Yucun and Yuexi catchments are from the Yangtze River Basin; Tunxi catchment is from Southeast River Basin; Hezikou catchment is from Pearl River Basin.

Table 4. Effect contributions of control factor categories on the temporal variabilities of flood event classes in the individual catchments

Classes	Catchments	River Basins	Meteorology	Land cover	Combination	Total
Class1	Hezikou	Pearl	49.4%	0.0%	17.3%	66.7%
	Liangshuikou	Yangtze	92.4%	3.8%	0.1%	96.3%
Class2	Shimenkan	Yangtze	87.1%	0.0%	3.6%	90.7%
	Tangdukou	Yangtze	95.9%	0.0%	0.0%	95.9%
	Xiaogulu	Pearl	71.9%	0.0%	24.9%	96.8%
Class3	-	-	-	-	-	-
Class4	Hezikou	Pearl	82.1%	0.0%	16.0%	98.1%
	Liangshuikou	Yangtze	76.8%	0.0%	10.2%	87.0%
Class5	-	-	-	-	-	-

4.4.2 Control factors and their contributions for the lumped analysis

The Monte Carlo permutation tests across the entire study area suggest that the meteorological category is the most important (Figure 8), particularly the precipitation amount and intensity, and the aridity index during the events with the correlation coefficients of 0.33–0.74, 0.20–0.38 and 0.29–0.41, respectively. The significant factor number in the catchment attribute category is less, which are mainly the mean catchment length, river density and ratio of river width to depth with the correlation coefficients of 0.18–0.32, 0.15–0.24 and 0.21–0.30, respectively. In the land cover category, only the grassland area ratio is significant in the Class 1 with the correlation coefficient of 0.21.

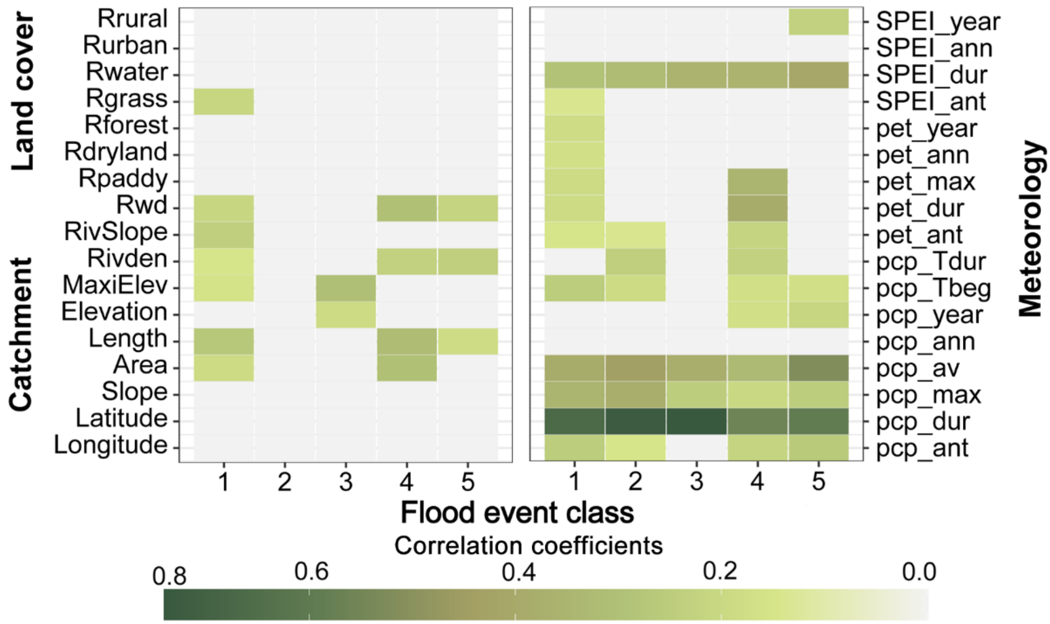


Figure 8. Significant control factors and their correlation coefficients for the variabilities of individual flood event classes (i.e., Classes 1–5). The gray color means the control factor without statistical significance.

In the Class 1, the significant control factors are the meteorological factors in the antecedent seven days ($r_{\mathrm{pcp\_ant}}$=0.25, $r_{\mathrm{pet\_ant}}$=0.15 and $r_{\mathrm{SPEI\_ant}}$=0.14), during the events ($r_{\mathrm{pcp\_dur}}$=0.67, $r_{\mathrm{pcp\_av}}$=0.39, $r_{\mathrm{pcp\_max}}$=0.35, $r_{\mathrm{pcp\_Tbeg}}$=0.25, $r_{\mathrm{pet\_dur}}$=0.19, $r_{\mathrm{pet\_max}}$=0.19 and $r_{\mathrm{SPEI\_dur}}$=0.29), and at the annual scale ($r_{\mathrm{pet\_ann}}$=0.17 and $r_{\mathrm{pet\_year}}$=0.18) in the meteorological category, the catchment area (r_{Area}=0.19), mean length (r_{Length}=0.27), catchment maximum elevation (r_{MaxiElev}=0.16), river density (r_{Rivden}=0.15) and slope (r_{RivSlope}=0.24) and ratio of river width to depth (r_{Rwd}=0.21) in the catchment attribute category, and the grassland area ratio (r_{Rglass}=0.21) in the land cover category. There are 72.7% of total spatial and temporal variabilities of flood events explained by all the control factor categories, in which 43.9% of total variabilities are explained by the meteorological category, followed by the combined impact (22.7%), catchment attribute category (4.2%) and land cover category (1.5%), respectively (Figure 9a).

The significant control factors are mainly the meteorological factors in the antecedent seven days and during the flood events for the Class 2, the meteorological factors during the flood events and catchment elevation for the Class 3, the meteorological factors in the antecedent seven days, during the flood events and at the annual scale, and the catchment factors related to slope and river for the Classes 4 and 5, respectively. The specific factors are the precipitation and potential evapotranspiration in the antecedent seven days (r_{pcp_ant}=0.15 and r_{pet_ant}=0.14), precipitation and aridity index during the flood events (r_{pcp_dur}=0.73, r_{pcp_av}=0.44, r_{pcp_max}=0.38, r_{pcp_Tbeg}=0.19, r_{pcp_Tdur}=0.24 and r_{SPEI_dur}=0.32) for the Class 2, the precipitation and aridity index during the flood events (r_{pcp_dur}=0.74, r_{pcp_av}=0.38, r_{pcp_max}=0.25, and r_{SPEI_dur}=0.36) in the meteorological category, and the catchment center elevation ($r_{Elevation}$=0.19) and maximum elevation ($r_{MaxiElev}$=0.31) in the catchment attribute category for the Class 3, the precipitation and potential evapotranspiration in the antecedent seven days (r_{pcp_ant}=0.22 and r_{pet_ant}=0.22), precipitation, potential evapotranspiration and aridity index during the events (r_{pcp_dur}=0.56, r_{pcp_av}=0.33, r_{pcp_max}=0.20, r_{pcp_Tbeg}=0.17, r_{pcp_Tdur}=0.23, r_{pet_dur}=0.39, r_{pet_max}=0.35, and r_{SPEI_dur}=0.36) and at the annual scale (r_{pcp_year}=0.17) for the meteorological attribute category, and the catchment area (r_{Area}=0.30), mean length (r_{Length}=0.32), river density (r_{Rivden}=0.23) and ratio of river width to depth (r_{Rwd}=0.30) in the catchment attribute category for the Class 4, and the precipitation in the antecedent seven days (r_{pcp_ant}=0.26), precipitation, potential evapotranspiration and aridity index during the events (r_{pcp_dur}=0.59, r_{pcp_av}=0.52, r_{pcp_max}=0.25, r_{pcp_Tbeg}=0.17 and r_{SPEI_dur}=0.41) and at the annual scale (r_{pcp_year}=0.21 and r_{SPEI_year}=0.23) for the meteorological attribute category, and the catchment mean length (r_{Length}=0.18), river density (r_{Rivden}=0.24) and ratio of river width to depth (r_{Rwd}=0.22) in the catchment attribute category for the Class 5, respectively.

For the contributions of individual control factor category, 73.3%, 85.4%, 65.9% and 65.7% of total spatial and temporal variabilities of flood events are explained by all the control factor categories in the Classes 2–5, respectively (Figure 9b–e). The meteorological category explains most of the variabilities, i.e., 46.6%, 50.5%, 39.2% and 36.5% in the Classes 2–5, respectively. The combined impact takes second place, which explains 22.8%, 33.0%, 20.6% and 23.7% of total variabilities in the Classes 2–5, respectively, followed by the catchment attribute category (i.e., 0.0%, 5.8%, 6.1% and 5.5% in the Classes 2–5, respectively). The impacts of land cover category in the Classes 2–5 are not significant.

Therefore, the total variabilities of flood events in the Class 1 are mainly controlled by the total precipitation amount and its intensity during the events which determine the magnitudes of total flood yield and flood peak, the catchment slope length and river slope which affect the flood routing processes, e.g., total duration of flood event and occurrence time of flood peak. The total variabilities in the Class 2 are also mainly controlled by the total precipitation amount and its intensity during the events. The total variabilities in the Class 3 are mainly controlled by the total precipitation amount, its intensity and the aridity index during the events which determine the total magnitudes and occurrence time of flood yield, and the catchment elevation which determine the flood routing time. The total variabilities in the Class 4 are mainly controlled by the total precipitation amount, potential evapotranspiration and the aridity index during the events which determine the total magnitude and occurrence time of flood yield, and evapotranspiration, as well as the catchment area, slope and river morphology which determine the flood routing time and river storage capacity. The total variabilities in the Class 5 are mainly controlled by the total precipitation amount and the aridity index during the events which determine the total magnitudes and occurrence time of flood yield, as well as the river density which determine the flood routing time in the river system.

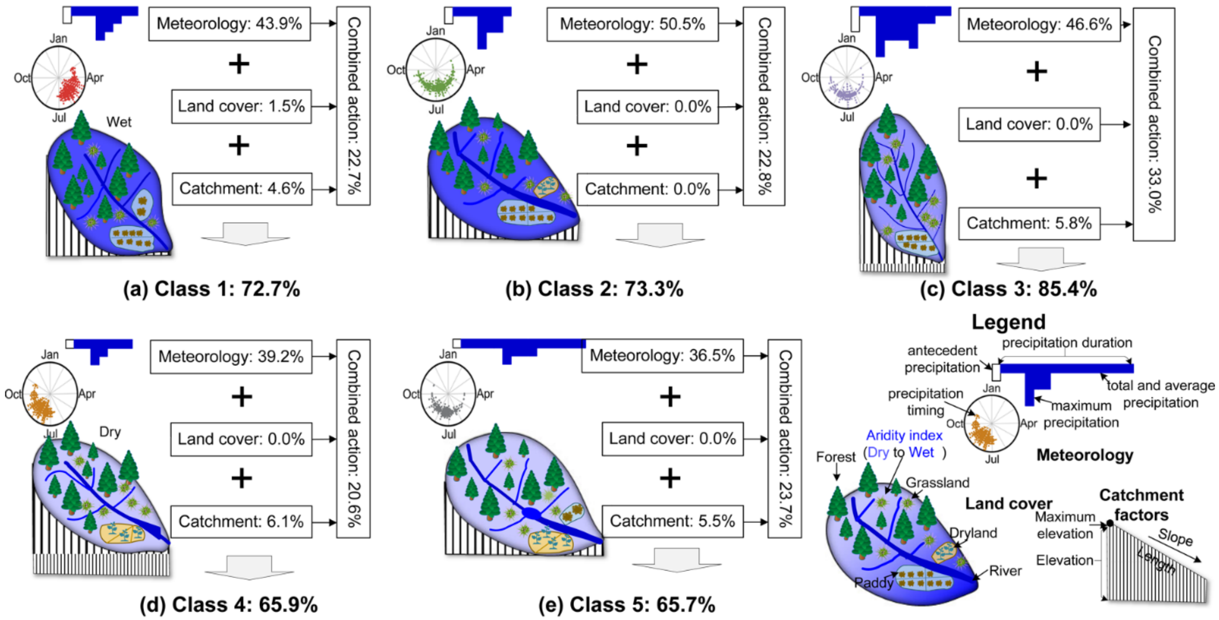


Figure 9. Contributions of control factor categories on the spatial and temporal variabilities of flood event classes 1–5 (a–e)

4.4.3 Control mechanisms in the individual flood event classes

In both the individual catchments and the entire region, the dominant control factors of all the flood event classes are the total and mean precipitation volumes, the maximum precipitation intensity, the aridity index and the precipitation timing during the events, the precipitation in the antecedent days in the meteorological category (Figures 10 and S6 in the Supplement). Therefore, the flood events in Class 1 are mainly caused by the rainfall with low volume and intensity before the wet season in the wet, steep and low-latitude catchments. The events in Class 2 are mainly caused by the short rainfall with high mean intensity in the wet low-latitude catchments. The events in Class 3 are mainly caused by the long rainfall with high volume and intensity in the small catchments of high altitude and low latitude. The events in Class 4 are mainly caused by the short rainfall with low volume and intensity in the latter wet season in the dry, steep and small catchments of high altitude and latitude. The events in Class 5 are mainly caused by the long rainfall with high volume and low mean intensity in the dry, gentle and large mid-latitude catchments.

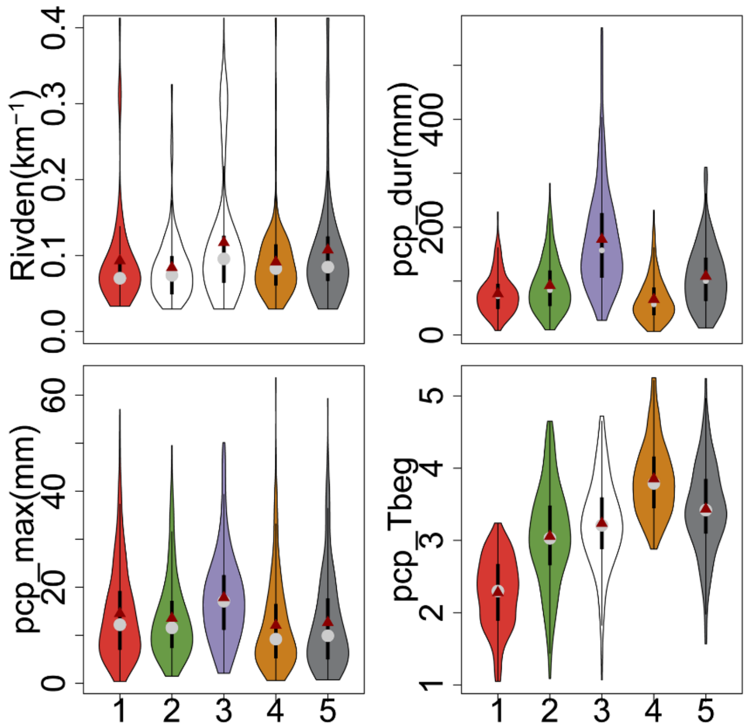


Figure 10. Variations of four critical control factors among Classes 1–5. The solid darkred dot and gray dot define the mean and 50th percentile values, respectively. Each black box means the 25th and 75th percentile values, and the vertical line defines the minimum and maximum values without outliers. The violin shape means the frequency distribution of control factor, and the unfilled shape means the control factor without statistical significance.

5. Discussion

Flood classification has strong advantages in systematically identifying manageable classes from a large number of historical flood events based on the similarity of flood response characteristics (Arthington *et al*., 2006; Kuentz *et al*., 2017; Poff *et al*., 2007; Sikorska *et al*., 2015; Sivakumar *et al*., 2015). Flood events in the same class are widely accepted to have similar hydrological responses caused by similar meteorological or underlying surface conditions (Sikorska *et al*., 2015). Therefore, it is more efficient to investigate flood event changes and their cause mechanisms in a comprehensive manner than individual event analyses (Zhang *et al*., 2012). It is expected to provide more useful flood response characteristics for flood disaster management purposes (e.g., early warning and quick design of flood control plans) and provide deep insights to investigate riverine ecological and environmental response mechanisms.

In our study, the flood event classes are identified based on the entire flood response characteristics, which cover not only the flood magnitude metrics (e.g., large, moderate and small floods) but also the event shape metrics (e.g., fast or slow floods).

Therefore, our study captures more detailed response dynamics of flood events than the predefined classes reported by several existing studies, such as flash floods, short-rain floods, rain-on-snow floods or snowmelt floods (Brunner *et al*., 2018; Merz and Blöschl, 2003; Sikorska *et al*., 2015). The specific values and boundaries of flood response metrics of individual classes were difficult to quantitatively compare with most existing studies because the adopted classification methods were usually different. However, the flood event classes with similar hydrographs or response mechanisms were also found in the existing studies. Classes 1 and 2 are mainly in the southern China, particularly in the Pearl and Yangtze River Basins, which are controlled by the temperate climate without a dry season. Storms with high intensities and short durations before the wet season in the southern China are likely to cause flood events with great magnitudes and variabilities (Class 1) or fast flood events with a high single peak and short durations (Class 2) (Gao *et al*., 2018). The flood response characteristics in these two classes are similar to the flash floods and short-rain floods in Austria (Merz and Blöschl 2003), and fast events in Switzerland (Brunner *et al*., 2018) and China (Zhai *et al*., 2021). Class 3 is mainly in the Southeast River Basin controlled by the tropical cyclone climate. Severe storms with high intensities and durations are likely to cause high slow flood events with multiple peaks (Class 3) (Yin *et al.*, 2010; Zhang *et al*., 2020). The flood response characteristics are similar to the high unit peak flood in the west coast of the USA (Saharia *et al*., 2017) because both the response characteristics were mainly controlled by subtropical or tropical storms near the ocean in the Cf climate type. They are also similar to the slow events in China (Zhai *et al*., 2021) because the rates of positive changes are 0.01–0.94 h^{-1} in our study and 0.04–1.78 h^{-1} in China (Zhai *et al*., 2021), and the rates of negative changes are 0.01–0.33 h^{-1} in our study and 0.02–0.25 h^{-1} in China (Zhai *et al*., 2021). Class 4 is mainly in the northern China controlled by the cold climate with dry winters. The heavy storms ahead of westerlies trough mainly occur in the latter wet season in this region, which usually have low intensities and short durations (Gao *et al*., 2018). Thus, they are likely to cause the small fast flood events (Class 4), whose mean flood peak magnitude and coefficients of variation are 0.47 $m^3/s/km^2$ and 0.86, respectively. The similar flood events are also reported, e.g., the low flashiness floods with the mean flood peak magnitude of 0.20–0.25 $m^3/s/km^2$ and the mean coefficients of variation of approximate 0.90 in the northern part of central–eastern Europe (Kuentz *et al*., 2017), which is also controlled by the similar climate type (i.e., Df). Class 5 is mainly in the south–north climate zone of China (i.e., Huaihe River Basin), which has the dual climate characteristics of both south and north monsoons. Storms characterized by a long period of continuous rainy meteorological with high frequency and low intensities (e.g., Meiyu rainfalls) in the earlier wet season are likely to cause moderate slow flood events with long durations (Gao et al., 2018; Sampe and Xie, 2010). The flood response characteristics are similar to the intermediate flood events in China (Zhai *et al*., 2021). For example, the coefficients of variation are 0.65–3.15 in our study and 0.78–3.07 in China (Zhai *et al*., 2021). The rates of positive and negative changes are 0.02–8.00 h^{-1} and 0.01–0.64 h^{-1} in our study, respectively, while those reported in Zhai *et al*. (2021) were 0.36–4.90 h^{-1} and 0.09–0.46 h^{-1} in China, respectively. Therefore, the classification is helpful to deeply investigate the control mechanisms of flood events, which is easy to transfer to predict flood events with similar control factors (Sikorska *et al*., 2015).

The meteorological, land cover and catchment attribute categories are mainly reported to affect the flood generation and routing processes, and could be widely-accepted as the critical control factors of spatial and temporal differences of flood event classes (Ali *et al*., 2012; Brunner *et al*., 2018; Merz and Blöschl, 2003; Zhang *et al*., 2022). Our results also find that the meteorological category is dominant, which explain 49.4–95.9% and 36.5–50.5% of the flood event differences in individual classes at catchment scale and in the entire region, respectively. Similar results were reported in Kuentz *et al*. (2017), which are that the climatic variables (e.g., precipitation, temperature and aridity index) play the most important role for 75% of total flow signatures and catchment attributes (e.g., area, elevation, slope and river density) are more important for the flood flashiness. The main significant meteorological factors are the precipitation volume, intensity and the aridity index during the events. The main explanation is that the precipitation and aridity index during the flood events directly affect the hydrograph through flood generation, e.g., total volume and peak, variability, duration, rate of changes and peak number (Merz and Blöschl, 2003; Aristeidis *et al*., 2010). Additionally, these control factors in the antecedent days directly affect the antecedent soil moisture, which determine the initial losses of precipitation and the runoff generation timing during the flood events (Hall and Blöschl, 2018; Xu *et al*., 2023). The contribution of meteorological category is the largest in the Class 2, particularly in the Tangdukou catchment of Yangtze River Basin because the flood events in this class usually show quick responses to the precipitation, while the contribution is the lowest in the Class 5 because the river density and river morphology play important roles in the flood storage capacity and routing time in the river system.

Secondly, the catchment attributes (e.g., geographical location and topography) mainly affect the hydrograph patterns through flood routing (Berger and Entekhabi, 2001; Ali *et al*., 2012), and the identified factors in our study are the catchment area and length, river density and ratio of river width to depth. For example, a catchment with longer routing length, larger routing area, river density and ratio of river width to depth usually has larger flood regulation and storage capacity, and thus generates the slow flood events, while a catchment with shorter routing length, smaller routing area, river density and ratio of river width to depth usually has weaker flood regulation and storage capacity, and thus generates the fast flood events (Zhang *et al*., 2020). However, the comprehensive contributions of catchment attributes are not considerable, i.e., only 0.0–6.1% in the entire region because the catchment attributes do not always well match the flood event responses (Kuentz *et al*., 2017; Ali *et al*., 2012). The contributions of catchment attribute category in the slow flood event classes (e.g., Classes 3 and 5) are usually larger than those in the fast flood event classes (e.g., Classes 1, 2 and 4) because the catchment attribute factors are significantly correlated with the flood response metrics in the Classes 3 and 5, particularly the catchment maximum elevation and river density. Furthermore, the location, annual precipitation, potential evapotranspiration and aridity index mainly affect the overall catchment hydrological conditions (Berger *et al*., 2001; Kennard *et al*., 2010). Finally, the land covers mainly determine the precipitation intercept and retention processes, which directly affect the flood variability and rate of changes (Kuentz *et al*., 2017; Merz *et al*., 2020). For example, catchments with greater vegetation covers (e.g., forest, grassland) usually generate the slow flood events, while catchments with weaker vegetation covers (e.g., rural and urban lands) usually generate the fast flood events (Kuentz *et al*., 2017; Zhai *et al*., 2021). However, all the catchments selected in our study are mainly in the river source regions with good vegetation coverages with mean area percentages of 67.0% for forest and 6.6% for grassland. The spatial and temporal differences of land covers are not remarkable so that it only explains 3.8% and 1.5% of the flood event differences in Class 1 at the Liangshuikou catchment of Yangtze River Basin and in the entire region, respectively.

Our study provided an approach to investigate some manageable flood event classes from massive events at large scale and to quantify the meteorological and physio-geographical controls of spatial and temporal variabilities of flood event classes. The approach could be applied easily to other regions or countries if a great number of flood events were collected. All the selected flood events were sufficient to represent the flood response characteristics of headstream catchments in main river basins of China. Thus, our classification results and the control mechanisms of variability of flood event classes would be applied in other regions with similar climate types. However, several works should be paid attention for further improvements of our study. Firstly, total flood event number is the main restricted factor for the classification performance, the flood event class representativeness and their control mechanisms at catchment scale (Merz and Blöschl, 2003; Olden *et al*., 2012; Sikorska *et al*., 2015; Tarasova *et al*., 2020). It could be overcome effectively by adopting large flood event numbers of individual classes (i.e., approximately 10% of total events at least in our study) for the classification (Zhang *et al*., 2020). However, not all the control mechanisms of flood event classes were well explained because of the insufficient flood events, which were mainly in the Songliao and Yellow River Basins, or most catchments expect the Shimenkan, Liangshuikou and Tangdukou catchments of the Yangtze River Basin, Xiaogulu and Hezikou catchments of the Pearl River Basin. The representatives of individual classes should be further investigated particularly in the basins with low densities of flood events. Secondly, the class boundaries of most flood response metrics were not clear using the inductive classification approaches (Parajka *et al*., 2005; Sikorska *et al*., 2015), e.g., the flood magnitude, rates of positive and negative changes in our study. Although the predefined the sharp thresholds of all the flood response metrics are beneficial to clearly separate the flood events using the classification tree methods (e.g., decision tree, crisp tree), the predefinition is still challenging (Sikorska *et al*., 2015; Brunner *et al*., 2017; Tarasova *et al*., 2020). Finally, the control mechanism deduction was mainly based on the statistical detection of control factors and their contributions. The combined impacts of different control factor categories were still difficult to be clearly explained using the adopted statistical analysis method (i.e., the constrained rank analysis in our study).

6. Conclusions

In our study, the main flood event classes characterized by multiple flood response metrics are identified in 68 headstream catchments using the hierarchical and partitional clustering methods. The control mechanisms of different flood event classes are investigated using the constrained rank analysis and Monte Carlo permutation test. Results are summarized as follows: the partitional clustering method (i.e., *k*-medoids) performs better than the hierarchical method, and the optimal five flood event classes are identified which are the moderately fast flood event class (Class 1), the highly fast flood event class (Class 2), the highly slow and multipeak flood event class (Class 3), the slightly fast flood event class (Class 4) and the moderately slow flood event class (Class 5). Most of the flood event differences among individual classes are explained by the meteorological, land cover and catchment attribute factors. The flood event differences in Class 3 (85.4%) are well explained, followed by Classes 2 (73.3%), 1 (72.7%), 4 (65.9%) and 5 (65.7%). The meteorological category is the most significant among all the control factors, particularly the precipitation factors (e.g., volume, intensity) and aridity index during the flood events.

This study preliminarily investigates the flood event classes in space and time in China, which is beneficial to explore the comprehensive formation mechanisms of flood events and the critical control factors, and provides the scientific foundation for flood event prediction and control. In future, more unimpaired flood events could be collected to strengthen the representativeness of flood event classes, and to further support the control mechanism analysis of flood classes at individual catchments. The combined impacts of control factor categories could also be further decomposed into the impacts of individual factors using the hydrological model with strong physical mechanism.

Acknowledgements

This study was supported by the National Natural Science Foundation of China (No. 42171047), the National Key Research and Development Program of China (No. 2021YFC3201102) and the CAS-CSIRO Partnership Joint Project of 2024 (No. 177GJHZ2023097MI). We extend our gratitude to the editors and anonymous reviewers for their insightful feedback and constructive comments.

Author contribution

Yongyong Zhang: Conceptualization, Methodology, Formal Analysis, Writing-Original draft preparation, Writing-Reviewing and Editing, Funding acquisition; **Yongqiang Zhang**: Conceptualization, Writing-Reviewing and Editing; **Xiaoyan Zhai:** Data curation, Formal Analysis, Writing-Reviewing and Editing, Funding acquisition; **Jun Xia:** Conceptualization, Writing-Reviewing and Editing; **Qiuhong Tang:** Conceptualization, Writing-Reviewing and Editing; **Wei Wang:** Data Processing, Formal Analysis; **Jian Wu:** Formal Analysis; **Xiaoyu Niu:** Formal Analysis; **Bing Han:** Formal Analysis

Code/Data availability

The geographic information system data sources are obtained from the Data Center of Resources and Environmental Science, Chinese Academy of Sciences (http://www.resdc.cn/). The historical flood events and synchronous precipitation are collected from the Hydrological Yearbooks of the Songliao, Yellow, Huaihe, Yangtze, Southeast and Pearl River Basins which are published by the Ministry of Water Resources of the People’s Republic of China (http://xxfb.mwr.cn/sq_djdh.html). The daily precipitation and temperature observations are collected from the China Meteorological Data Sharing Service System (http://cdc.cma.gov.cn/home.do).

Competing interests

The authors declare no conflicts of interest relevant to this study.

References

Ali, G., Tetzlaff, D., Soulsby, C., McDonnell, J. J., and Capell, R.: A comparison of similarity indices for catchment classification using a cross-regional dataset, Adv. Water Resour., 40, 11–22, https://doi.org/10.1016/j.advwatres.2012.01.008, 2012.

Antoine, L. and Sylvain J.: Using amap and ctc Packages for Huge Clustering. R News, 6(5):58-60, 2006.

Aristeidis, G. K., Tsanis, I. K., and Daliakopoulos, I. N.: Seasonality of floods and their hydrometeorologic characteristics in the island of Crete, J. Hydrol., 394(1–2), 90-100, https://doi.org/10.1016/j.jhydrol.2010.04.025, 2010

Arthington, A. H., Bunn, S. E., Poff, N. L., and Naiman, R. J.: The challenge of providing environmental flow rules to sustain river ecosystems, Ecol. Appl., 16, 1311–1318, https://doi.org/10.1890/1051-0761(2006)016[1311:TCOPEF]2.0.CO;2, 2006.

Black, A.R., Werritty, A: Seasonality of flooding: a case study of North Britain, J. Hydrol.,195:1–25, https://doi.org/10.1016/S0022-1694(96)03264-7, 1997.

Berger, K. P., and Entekhabi, D.: Basin hydrologic response relations to distributed physiographic descriptors and climate, J. Hydrol., 247(3-4), 169-182, https://doi.org/10.1016/S0022-1694(01)00383-3, 2001.

Berghuijs, W. R., Woods R. A., Hutton C. J., and Sivapalan M.: Dominant flood generating mechanisms across the United States, Geophys. Res. Letters, 43, 4382–4390, http://doi.org/10.1002/2016GL068070, 2016.

Brunner, M. I., Viviroli, D., Furrer, R., Seibert, J., and Favre, A. C.: Identification of flood reactivity regions via the functional clustering of hydrographs, Water Resour. Res., 54, 1852-1867, https://doi.org/10.1002/2017WR021650, 2018

Brunner, M. I., Viviroli, D., Sikorska, A. E., Vannier, O., Favre, A.C., and Seibert, J.: Flood type specific construction of synthetic design hydrographs, Water Resour. Res., 53,1390–1406, https://doi.org/10.1002/2016WR019535, 2017.

Charrad, M., Ghazzali, N., Boiteau, V., and Niknafs, A.: NbClust: An R package for determining the relevant number of clusters in a data set, J. Stat. Softw., 61(6), 1-36, http://hdl.handle.net/10.18637/jss.v061.i06,2014.

China Institute of Water Resources and Hydropower Research, and Research Center on Flood and Drought Disaster Prevention and Reduction, the Ministry of Water Resources. Atlas of Flash Flood Disasters in China. Sinomap Press, 2021. (in Chinese)

Dhakal, N., Jain, S., Gray, A. , Dandy, M., and Stancioff, E.: Nonstationarity in seasonality of extreme precipitation: a nonparametric circular statistical approach and its application, Water Resour. Res., 51(6), 4499-4515, https://doi.org/10.1002/2014WR016399, 2015.

Dong, Q., Chen, X., and Chen, T.: 2011. Characteristics and changes of extreme precipitation in the Yellow-Huaihe and Yangtze-Huaihe Rivers Basins, China, J. Climate, 24(14), 3781-3795, https://doi.org/10.1175/2010JCLI3653.1, 2011.

Fisher, N.I.: Statistical Analysis of Circular Data. Cambridge University Press, Cambridge, UK, 1993.

Gao, S.T., Zhou, Y.S., Ran, L.K.: A review on the formation mechanisms and forecast methods for torrential rain in China, Chinese Journal of Atmospheric Sciences, 42 (4), 833-846, https://doi.org/10.3878/j.issn.1006-9895.1802.17277, 2018. (in Chinese)

Hall, J., and Blöschl, G.: Spatial patterns and characteristics of flood seasonality in Europe, Hydrol. Earth Syst. Sc., 22, 3883–3901, https://doi.org/10.5194/hess-22-3883-2018, 2018.

Hargreaves, G. H., and Samani, Z. A.: Estimating potential evapotranspiration, J. Irrig. Drain. Div., 108(3), 225-230, https://doi.org/10.1016/0022-1694(82)90165-2,1982.

Hartigan, J. A., and Wong, M. A.: Algorithm AS 136: A K-means clustering algorithm, Appl. Stat., 28, 100–108, https://doi.org/10.2307/2346830, 1979.

Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., Kim, H., and Kanae, S.: Global flood risk under climate change, Nat. Clim. Change, 3, 816–821, https://doi.org/10.1038/NCLIMATE1911, 2013.

Jin, H., Chen, X., and Adamowski, J. H. S.: Determination of duration, threshold and spatiotemporal distribution of extreme continuous precipitation in nine major river basins in China, Atmos Res, 300, 107217, https://doi.org/10.1016/j.atmosres.2023.107217, 2024.

Jongman, B. Winsemius, H. C., Aerts, J. C. J. H., Coughlan de Perez, E., van Aalst, M. K., Kron, W., and Ward, P. J.: Declining vulnerability to river floods and the global benefits of adaptation, P. Natl. Acad. Sci. USA, 112, E2271–E2280, https://doi.org/10.1073/pnas.1414439112, 2015.

Kaufman, L. and Rousseeuw, P.J.: Finding Groups in Data: An Introduction to Cluster Analysis. Wiley, New York,1990.

Kennard, M. J., Pusey, B. J., Olden, J. D., Mackay, S. J., Stein, J. L., and Marsh, N.: Classification of natural flow regimes in Australia to support environmental flow management, Freshwater Biol., 55, 171–193, https://doi.org/10.1111/j.1365-2427.2009.02307.x, 2010

Kuentz, A., Arheimer, B., Hundecha, Y., and Wagener, T.: Understanding hydrologic variability across Europe through catchment classification, Hydrol. Earth Syst. Sc., 21, 2863–2879. https://doi.org/10.5194/hess-21-2863-2017, 2017.

Legendre, P., and Anderson, M. J.: Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments, Ecol. Monogr., 69(1), 1-24, https://doi.org/10.1890/0012-9615(1999)069[0001:DBRATM]2.0.CO;2, 1999.

Liu, J.Y., Feng, S.Y., Gu, X.H., Zhang, Y.Q., Beck, H.E., Zhang, J.W., and Yan, S.: Global changes in floods and their drivers, J. Hydrol., 614, 128553, https://doi.org/10.1016/j.jhydrol.2022.128553, 2020.

Mardia, K. V., Kent, J. T., and Bibby, J. M.: Multivariate Analysis, London, Academic Press, 1979.

Merz, R., and Blöschl, G.: A process typology of regional floods, Water Resour. Res., 39(12), 1340, https://doi.org/10.1029/2002WR001952, 2003.

Merz, R., Tarasova, L., and Basso, S.: The Flood Cooking Book: ingredients and regional flavors of floods across Germany, Environ. Res. Lett., 15(11), 114024, https://doi.org/10.1088/1748-9326/abb9dd, 2020.

Ministry of Water Resources of the People's Republic of China: Bulletin of flood and drought disasters in China from 2009 to 2018. China Water & Power Press. Retrieved from http://www.mwr.gov.cn/sj/tjgb/zgshzhgb/, 2020a. (in Chinese)

Ministry of Water Resources of the People's Republic of China: Code for hydrologic data processing (SL/T 247—2020), 2020b. (in Chinese)

Olden, J. D., Kennard, M. J., and Pusey, B. J.: A framework for hydrologic classification with a review of methodologies and applications in ecohydrology, Ecohydrology, 5, 503–518, https://doi.org/10.1002/eco.251, 2012.

Parajka, J., Merz, R., and Blöschl, G.: A comparison of regionalization methods for catchment model parameters, Hydrol. Earth Syst. Sc., 9, 157–171, https://doi.org/10.5194/hess-9-157-2005, 2005.

Peel, M. C., Finlayson, B. L., and McMahon, T. A.: Updated world map of the Köppen-Geiger climate classification, Hydrol. Earth Syst. Sc., 11, 1633–1644, https://doi.org/10.5194/hess-11-1633-2007, 2007.

Poff, N.L., Allan, J.D., Bain, M.B., Karr, J.R., Prestegaard, K.L., Richter, B.D., Sparks, R.E., and Stromberg, J.C.: The natural flow regime: a paradigm for river conservation and restoration, Bioscience, 47, 769–784, 10.1080/09647778409514905, 1997.

Poff, N. L., and Zimmerman, J. K. H.: Ecological responses to altered flow regimes: a literature review to inform environmental flows science and management, Freshwater Biol., 55, 194–205, https://doi.org/10.1080/09647778409514905, 2010.

Poff, N. L., Olden, J. D., Merritt, D., and Pepin, D.: Homogenization of regional river dynamics by dams and global biodiversity implications, P. Natl. Acad. Sci. USA, 104, 5732–5737, https://doi.org/10.1073/pnas.0609812104, 2007.

R Development Core Team: R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org, 2010.

Richter, B.D., Baumgartner, J.V., Powell, J., Braun, D.P.: A method for assessing hydrological alteration within ecosystems, Conserv. Biol., 10(4): 1163–1174, https://doi.org/10.1046/j.1523-1739.1996.10041163.x, 1996.

Rustomji, P., Bennett, N., and Chiew, F.: Flood variability east of Australia's great dividing range, J. Hydrol., 374(3-4), 196-208, https://doi.org/10.1016/j.jhydrol.2009.06.017, 2009.

Saharia, M., Kirstetter, P. E., Vergara, H., Gourley, J. J., and Hong, Y.: Characterization of floods in the United States, J. Hydrol., 548, 524-535, https://doi.org/10.1016/j.jhydrol.2017.03.010, 2017.

Sampe, T., and Xie, P.: Large-scale dynamics of the Meiyu baiu rainband: Environmental forcing by the westerly jet. J. Climate, 23, 113–133, https://doi.org/10.1175/2009JCLI3128.1, 2010.

Shi, P., Zhang, Y., Li, Z., Li, P., and Xu, G.; Influence of land use and land cover patterns on seasonal water quality at multi-spatial scales, Catena, 151, 182-190, https://doi.org/10.1016/j.catena.2016.12.017, 2017.

Sikorska, A. E., Viviroli, D. and Seibert, J.: Flood-type classification in mountainous catchments using crisp and fuzzy decision trees, Water Resour. Res., 51, 7959–7976, https://doi.org/10.1002/2015WR017326, 2015.

Sivakumar, B., Singh, V., Berndtsson, R., and Khan, S.: Catchment classification framework in hydrology: challenges and directions, J. Hydrol. Eng., 20, A4014002. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000837, 2015

Tarasova, L., Basso, S., Wendi, D., Viglione, A., Kumar, R., and Merz, R.: A process-based framework to characterize and classify runoff events: The event typology of Germany, Water Resour. Res., 56, e2019WR026951, https://doi.org/10.1029/2019WR026951, 2020.

Tarasova, L., Basso, S., Zink, M. and Merz, R.: Exploring controls on rainfall-runoff events: 1. Time series-based event separation and temporal dynamics of event runoff response in Germany, Water Resour. Res. 54, 7711–7732, https://doi.org/10.1029/2018WR022587, 2018.

Tarasova, L., Merz, R., Kiss, A., Basso, S., Günter, B., Merz, B., Viglione, A., Plötner, S. Guse, B., Schumann, A., Fischer, S., Ahrens, B., Anwar, F., Bárdossy, A., Bühler, P., Haberlandt, U., Kreibich, H., Krug, A., Lun, D., Müller-Thomy, H., Pidoto, R., Primo, C., Seidel. J., Vorogushyn, S., Wietzke, L.: Causative classification of river flood events, WIRES Water, 6(4), e1353, https://doi.org/10.1002/wat2.1353, 2019.

ter Braak, C. J. F.: Canonical Correspondence Analysis: a new eigenvector technique for multivariate direct gradient analysis, Ecology, 67, 1167-1179, https://doi.org/10.2307/1938672, 1986.

Villarini, G. On the seasonality of flooding across the continental United States, Adv. Water Resour., 87, 80-91, https://doi.org/10.1016/j.advwatres.2015.11.009, 2016.

Wang, H., Liu, J.G., Klaar, M., Chen, A.F., Gudmundsson, L., and Holden, J.: Anthropogenic climate change has influenced global river flow seasonality, Science, 383(6686), 1009-1014, https://doi.org/10.1126/science.adi9501, 2024.

Xu, Z., Zhang, Y., Blöschl, G., and Piao, S.: Mega forest fires intensify flood magnitudes in southeast Australia, Geophys. Res. Lett., 50, e2023GL103812, https://doi.org/10.1029/2023GL103812, 2023.

Yin, Y.Z., Gemmer, M., Luo, Y., Wang, Y.: Tropical cyclones and heavy rainfall in Fujian Province, China. Quatern int., 226(1–2):122-128, https://doi.org/10.1016/j.quaint.2010.03.015, 2010.

Zhai, X. Y., Guo, L., and Zhang, Y. Y.: Flash flood type identification and simulation based on flash flood behavior indices in China, Sci. China Earth Sci., 64(7), 1140–1154, https://doi.org/10.1007/s11430-020-9727-1,2021.

Zhang, S.L., Zhou, L.M., Zhang, L., Yang, Y.T., Wei, Z.W., Zhou, S., Yang, D.W., Yang, X. F., Wu, X.C., Zhang, Y.Q., and Dai, Y.J.: Reconciling disagreement on global river flood changes in a warming climate, Nat. Clim. Change, 12, 1160–1167, https://doi.org/10.1038/s41558-022-01539-7, 2022.

Zhang, Y. Y., Arthington, A. H., Bunn, S. E., Mackay, S., Xia, J., and Kennard, M.: Classification of flow regimes for environmental flow assessment in regulated rivers: the Huai River Basin, China, River Res. Appl., 28, 989–1005, https://doi.org/10.1002/rra.1483, 2012.

Zhang, Y., Zhou, Y., Shao, Q., Liu, H., Lei, Q., Zhai, X., and Wang, X.: Diffuse nutrient losses and the impact factors determining their regional differences in four catchments from North to South China, J. Hydrol., 543, 577-594, https://doi.org/10.1016/j.jhydrol.2016.10.031,2016.

Zhang, Y. Y., Chen, Q. T., and Xia, J.: Investigation on flood event variations at space and time scales in the Huai River Basin of China using flood behavior classification, J. Geogr. Sci, 30(12), 2053-2075, https://doi.org/10.1007/s11442-020-1827-3, 2020.

Zhang, Y. Y., Zhang, Y. Q., Zhai, X., Xia, J., Tang, Q., Zhao, T., and Wang, W.: Predicting flood event class using a novel class membership function and hydrological modelling, Earth's Future, 12, e2023EF004081, https://doi.org/10.1029/2023EF004081,2024.