Supplement of Hydrol. Earth Syst. Sci., 29, 3257–3275, 2025
https://doi.org/10.5194/hess-29-3257-2025-supplement

Hydrology and Earth System Sciences
Open Access
EGU

Supplement of

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

Yongyong Zhang et al.

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

Text S1

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 drought index during the event with the correlation coefficients of 0.61–0.99, 0.58–0.99 and 0.50–0.98, respectively (Table 4 and Figure S2). 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 drought index in the corresponding year with the correlation coefficients of 0.53–1.00 and 0.45–0.93, respectively (Figure S4). 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 (Figures S3 and S5). 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 drought 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).

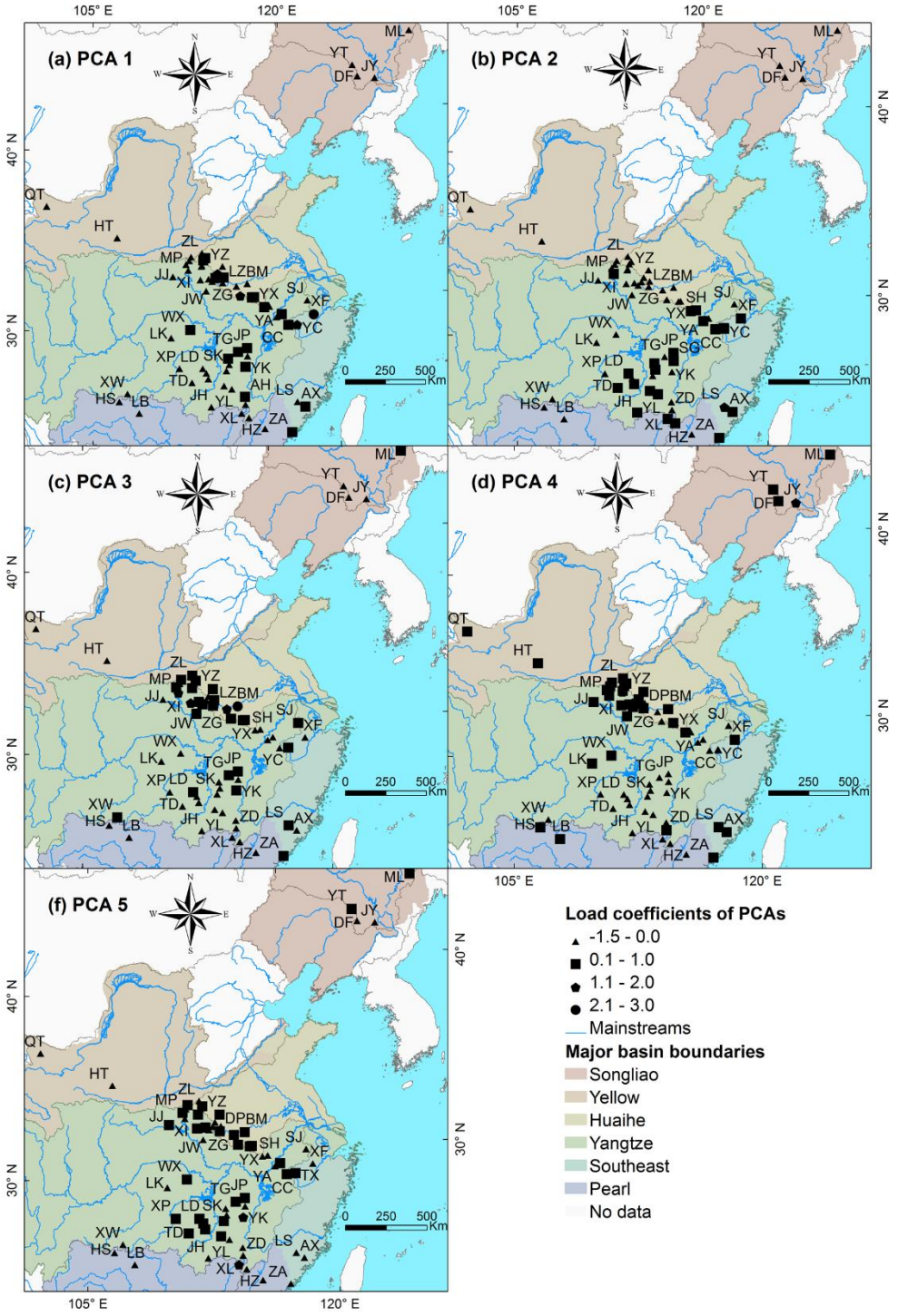

Figure S1. Spatial distributions of load coefficients of all the principal components.

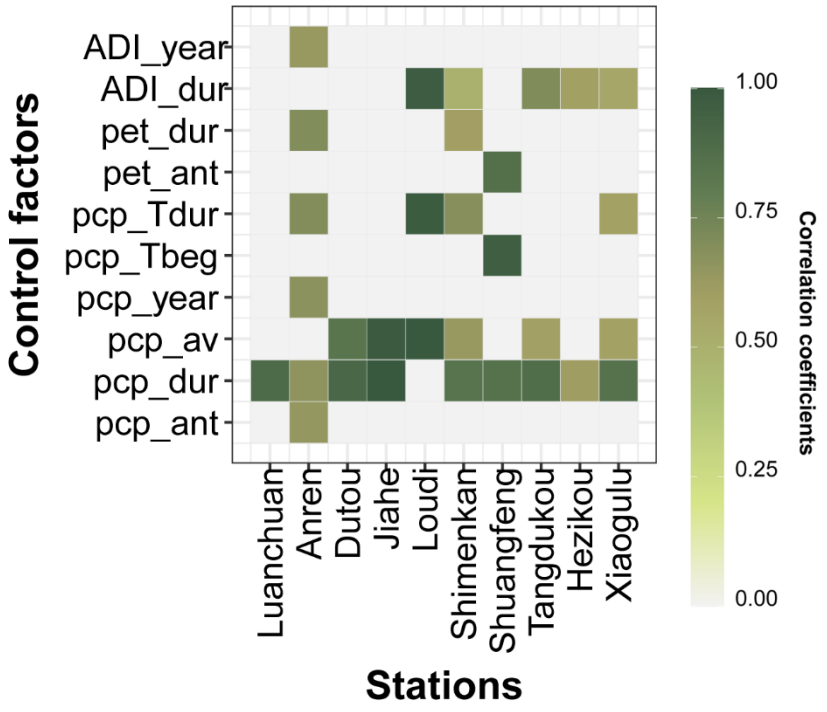

Figure S2. Significant control factors and their correlation coefficients for the temporal variabilities of flood event Class 2 in the individual catchments. The gray color means the control factor without statistical significance.

Note: Anren, Dutou, Jiahe, Loudi, Shimenkan, Shuangfeng and Tangdukou catchments are in the Yangtze River Basin; Luanchuan catchment is in the Yellow River Basin; Hezikou and Xiaogulu catchments are in the Pearl River Basin

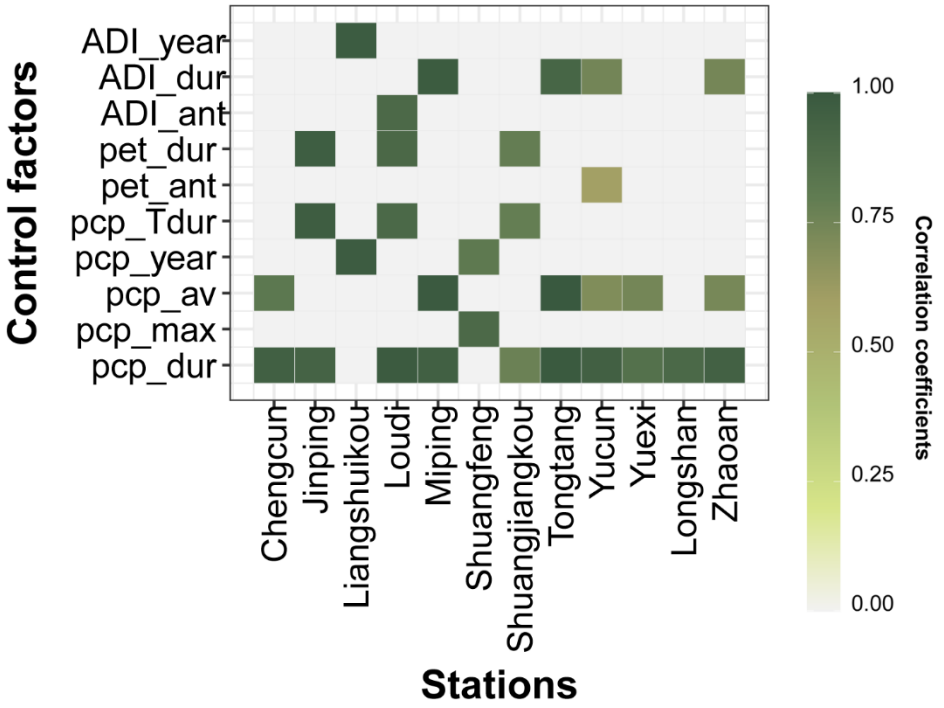

Figure S3. Significant control factors and their correlation coefficients for the temporal variabilities of flood event Class 3 in the individual catchments. The gray color means the control factor without statistical significance.

Note: Chengcun, Jinping, Liangshuikou, Loudi, Miping, Shuangfeng, Shuangjiangkou, Tongtang, Yucun and Yuexi catchments are in the Yangtze River Basin; Longshan and Zhaoan catchments are in the Pearl River Basin

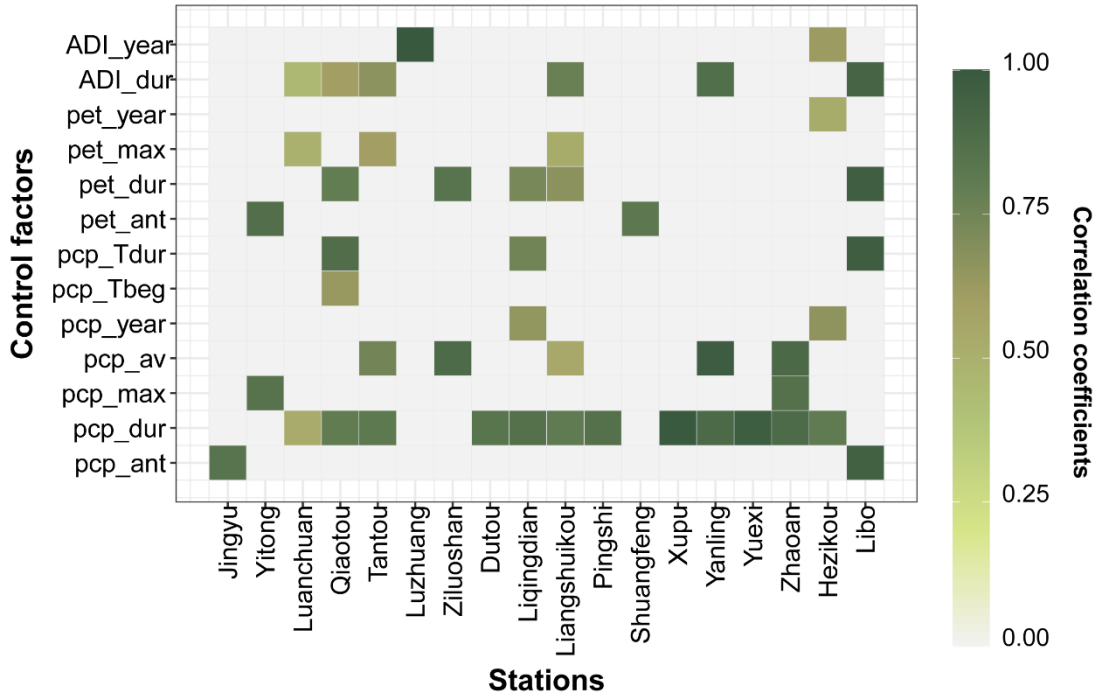

Figure S4. Significant control factors and their correlation coefficients for the temporal variabilities of flood event Class 4 in the individual catchments. The gray color means the control factor without statistical significance.

Note: Jingyu and Yitong catchments are in the Songliao River Basin; Luanchuan, Qiaotou and Tantou catchments are in the Yellow River Basin; Luzhuang and Ziluoshan catchments are in the Huaihe River Basin; Dutou, Liqingdian, Liangshuikou, Pingshi, Shuangfeng, Xupu, Yanling and Yuexi catchment are in the Yangtze River Basin; Zhaoan catchment is in the Southeast River Basin; Hezikou and Libo catchments are in the Pearl River Basin

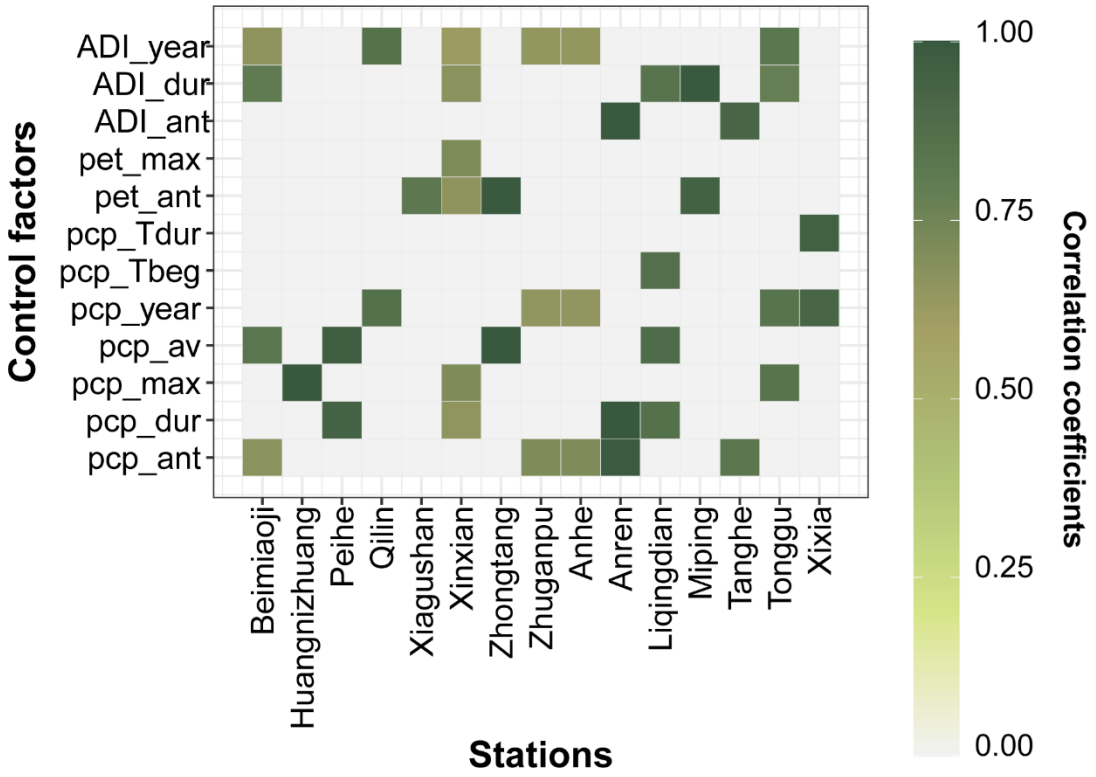

Figure S5. Significant control factors and their correlation coefficients for the temporal variabilities of flood event Class 5 in the individual catchments. The gray color means the control factor without statistical significance.

Note: Beimiaoji, Huangnizhuang, Peihe, Qilin, Xiagushan, Xinxian, Zhongtang and Zhuganpu catchments are in the Huaihe River Basin; Anhe, Anren, Liqingdian, Miping, Tanghe, Tonggu and Xixia catchments are in the Yangtze River Basin.

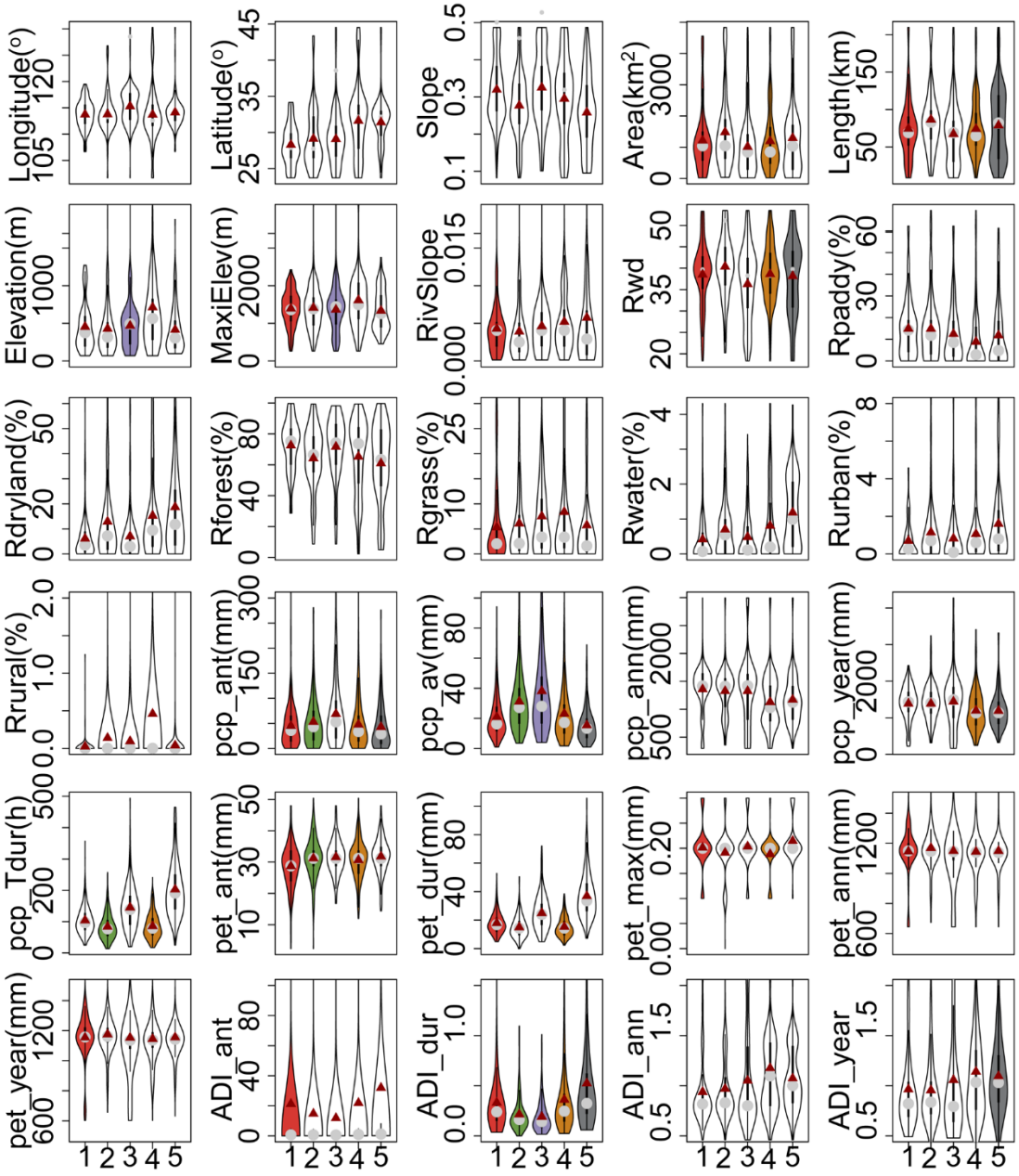

Figure S6. Variations of the other 30 critical control factors among Classes 1-5. The solid darkred dot and gray dot define the mean and 50th percentile values, respectively. Each black box means the 25th and 75th percentile values, and the vertical line defines the minimum and maximum values without outliers. The violin shape means the frequency distribution of control factor, and the unfilled shape means the control factor without statistical significance.

Table S1. Total numbers and densities of hydrological stations and flood events in different river basins

Basin	Area (10^4km^2)	Number		Density	
		Station	Flood event	Station (10^{-4} station/km^2)	Event (10^{-4} event/km^2)
Songliao River Basin	124.92	4	53	0.03	0.42
Yellow River Basin	75.24	4	104	0.05	1.38
Huaihe River Basin	27.00	13	215	0.48	7.96
Yangtze River Basin	180.85	38	844	0.21	4.67
Southeast River Basin	24.02	5	90	0.21	3.75
Pearl River Basin	45.36	4	140	0.09	3.09

Table S2. Criteria of classification performance assessment

ID	Criteria name	Abbreviation	Equation	Reference
1	Krzanowski-Lai	KL	$KL(q) = \left\lvert \frac{DIEF_q}{DIEF_{q+1}} \right\rvert$	Krzanowski and Lai 1988
2	Calinski-Harabasz	CH	$CH(q) = \frac{\mathrm{trace}(B_q)/(q-1)}{\mathrm{trace}(W_q)/(n-q)}$	Calinski and Harabasz 1974
3	Hartigan	Hartigan	$Hartigan = \left(\frac{\mathrm{trace}(W_q)}{\mathrm{trace}(W_{q+1})} - 1 \right)(n-q-1)$	Hartigan 1975
4	Cubic Clustering Criterion	CCC	$CCC = \ln\left[\frac{1-E(R^2)}{1-R^2} \right] \frac{\sqrt{\frac{np^*}{2}}}{(0.001+E(R^2))^{1.2}}$	Sarle 1983
5	Scott	Scott	$Scott = n\log\frac{\det(T)}{\det(W_q)}$	Scott and Symons 1971
6	Marriot	Marriot	$Marriot = q^2\det(W_q)$	Marriot 1971
7	Trcovw	TrCovW	$Trcovw = \mathrm{trace}(COV(W_q))$	Milligan and Cooper 1985
8	Tracew	TraceW	$Tracew = \mathrm{trace}(W_q)$	Milligan and Cooper 1985
9	Friedman	Friedman	$Friedman = \mathrm{trace}(W_q^{-1}B_q)$	Friedman and Rubin 1967
10	Silhouette	Silhouette	$Silhouette = \frac{\sum_{i=1}^{n} S(i)}{n}, Silhouette \in [-1,1]$	Rousseeuw 1987
11	Ratkowsky-Lance	Ratkowsky	$Ratkowsky = \frac{\bar{S}}{q^{1/2}}$	Ratkowsky and Lance 1978
12	Ball	Ball	$Ball = \frac{W_q}{q}$	Ball and Hall 1965
13	Ptbiserial	Ptbiserial	$Ptbiserial = \frac{[\bar{S}_b - \bar{S}_w][N_w N_b / N_t^2]^{1/2}}{s_d}$	Milligan 1980, 1981
14	Dunn	Dunn	$Dunn = \frac{\min_{1\le i<j\le q}(C_i, C_j)}{\max_{1\le k\le q}\mathrm{diam}(C_k)}$	Dunn 1974
15	Rubin	Rubin	$Rubin = \frac{\det(T)}{\det(W_q)}$	Friedman and Rubin 1967
16	C-Index	Cindex	$Cindex = \frac{S_w - S_{\min}}{S_{\max} - S_{\min}}, S_{\min} \neq S_{\max}, Cindex \in (0,1)$	Hubert and Levin 1976
17	Davies-Bouldin	DB	$DB(q) = \frac{1}{q}\sum_{k=1}^{q}\max_{k\neq l}\left(\frac{\delta_k + \delta_l}{d_{kl}}\right)$	Davies and Bouldin 1979
18	Duda	Duda	$Duda \geq 1 - \frac{2}{\pi p} - \sqrt{\frac{2(1-\frac{8}{\pi^2 p})}{n_m p}} = critValue\_Duda$	Duda and Hart 1973
19	Pseudo t^2	Pseudot2	$Pseudot2 = \frac{V_{kl}}{\frac{W_k + W_l}{n_k + n_k - 2}}$	Duda and Hart 1973
20	McClain-Rao	McClain	$McClain = \frac{\bar{S}_w}{\bar{S}_b} = \frac{S_w / N_w}{S_b / N_b}$	McClain and Rao 1975
21	SD validity	SDindex	$SDindex(q) = \alpha Scat(q) + Dis(q)$	Halkidi et al. 2000
22	SDbw validity	SDbw	$SDbw(q) = Scat(q) + Density.bw(q)$	Halkidi and Vazirgiannis 2001

Note: q is the number of clusters; n is the number of observations; p is the number of variables; B_q is the between-group dispersion matrix for data clustered into q clusters; W_q is the within-group dispersion matrix for data clustered into q clusters; R^2 is the coefficient of determination; T is the total sum of squares; S_b is the sum of the between-cluster distances; S_w is the sum of the within-cluster distances; $\bar{S}_b$ is the ratio of the S_b and N_b; $\bar{S}_w$ is the ratio of the S_w and N_w; N_w is the total number of pairs of observations belonging to the same cluster; N_b is the total number of pairs of observations belonging to different clusters; N_t is the total number of pairs of observations in the data set; S_{max} is the sum of the N_w largest distances between all the pairs of points in the entire data set; S_{min} is the sum of the N_w smallest distances between all the pairs of points in the entire data set (there are N_t such pairs); S_d is the standard deviation of all distances; $\bar{S}$ is the average of the ratios of sum and total sum of squares between the clusters for each variable; i is the number ranges from 1 to n; j is the number ranges from 1 to p; k, l and m is the cluster number ranges from 1 to q; C_i; C_j and C_k are the different clusters; n_k, n_l and n_m are the number of objects in cluster C_k, C_l and C_m, respectively; W_k, W_l and W_m are the squared errors of the different clusters; V_{kl} equals W_m minus W_k and then minus W_l; d_{kl} is the distance between centroids of clusters C_k and C_l; δ_k and δ_l are the standard deviation of the distance of objects in cluster C_k and C_l, respectively.

Table S3. Results of independence and linear correlation tests among different flood response metrics

Methods		Correlation coefficient								
		R	Q_{pk}	CV	T_{bgn}	T_{pk}	T_{drn}	RQ_r	RQ_d	N_{pk}
p-value for ANOVA test	R		**0.68**	**0.14**	*0.00*	**0.06**	**0.14**	**0.26**	**0.34**	**0.34**
	Q_{pk}	**0.00**		**0.41**	*0.02*	*-0.03*	**-0.18**	**0.75**	**0.77**	**0.08**
	CV	**0.00**	**0.00**		**0.06**	**-0.24**	**0.18**	**0.38**	**0.19**	**-0.21**
	T_{bgn}	*0.93*	*0.45*	**0.02**		**-0.12**	**0.07**	*0.04*	*0.04*	*-0.04*
	T_{pk}	**0.02**	*0.19*	**0.00**	**0.00**		**-0.14**	**-0.19**	**0.11**	**0.14**
	T_{drn}	**0.00**	**0.00**	**0.00**	**0.01**	**0.00**		**-0.19**	**-0.28**	**0.23**
	RQ_r	**0.00**	**0.00**	**0.00**	*0.12*	**0.00**	**0.00**		**0.68**	**-0.03**
	RQ_d	**0.00**	**0.00**	**0.00**	*0.15*	**0.00**	**0.00**	**0.00**		*0.02*
	N_{pk}	**0.00**	**0.00**	**0.00**	*0.17*	**0.00**	**0.00**	*0.31*	*0.38*	

Note: the bold value indicates that the test passes the 95% significance test, and the italic value indicates that the test does not pass the 95% significance test."

Table S4. Average, standard deviation, median, maximum and minimum of flood response metrics in different classes

Characteristic value	Class	R(mm·day^{-1})	Q_{pk}(mm·day^{-1})	CV	T_{bgn}	T_{pk}(%)	T_{drn}(h)	RQ_r(h^{-1})	RQ_d(h^{-1})	N_{pk}
Average± Standard Deviation	1	43.97±29.94	2.04±2.51	0.90±0.26	2.28±0.49	27.14±9.60	103.92±43.39	0.13±0.32	0.04±0.07	1.31±0.51
	2	45.81±34.01	2.21±2.52	0.87±0.25	3.06±0.69	50.64±10.28	83.82±41.20	0.08±0.14	0.08±0.12	1.32±0.50
	3	143.97±108.33	5.23±6.04	0.84±0.22	3.24±0.61	33.90±15.02	145.26±68.99	0.25±0.62	0.12±0.28	2.67±0.76
	4	33.31±26.64	1.69±2.11	0.86±0.26	3.85±0.51	26.11±9.09	85.73±39.97	0.14±0.30	0.04±0.08	1.24±0.43
	5	65.79±43.80	2.98±3.68	1.40±0.43	3.43±0.61	23.74±13.60	202.88±85.42	0.18±0.62	0.03±0.04	1.24±0.46
Median	1	35.63	1.17	0.89	2.30	27.27	97.01	0.05	0.02	1.00
	2	37.84	1.36	0.84	3.03	49.04	76.99	0.04	0.04	1.00
	3	115.53	3.09	0.82	3.21	32.09	139.01	0.07	0.03	3.00
	4	25.09	1.00	0.83	3.79	26.39	79.01	0.05	0.02	1.00
	5	57.11	1.92	1.32	3.42	21.26	190.99	0.04	0.01	1.00
Maximum	1	171.48	22.92	1.97	3.24	57.14	357.00	4.58	0.74	3.00
	2	194.87	19.84	1.81	4.65	86.96	256.99	1.24	1.06	3.00
	3	610.70	34.79	1.45	4.72	79.91	493.99	6.89	2.45	4.00
	4	174.43	21.02	2.12	5.25	55.67	241.01	3.50	0.91	3.00
	5	201.00	27.18	3.15	5.24	81.56	465.00	6.76	0.31	3.00
Minimum	1	3.22	0.13	0.33	1.05	4.17	25.01	0.00	0.00	1.00
	2	1.11	0.07	0.32	1.09	32.65	13.99	0.00	0.00	1.00
	3	7.79	0.14	0.32	1.07	4.47	19.99	0.00	0.00	1.00
	4	1.17	0.04	0.29	2.88	5.56	16.99	0.00	0.00	1.00
	5	1.54	0.07	0.65	1.57	1.61	25.01	0.00	0.00	1.00

Table S5. Flood event number and their percentages of individual classes in all the selected catchments

Basins	Stations	Abbreviations	Flood event number of class						Percentage(%)				
			1	2	3	4	5	Total	1	2	3	4	5
Songliao	Dongfeng	DF	0	3	1	9	1	14	0.0	21.4	7.1	64.3	7.1
	Jingyu	JY	0	3	1	9	0	13	0.0	23.1	7.7	69.2	0.0
	Muling	ML	0	0	2	7	3	12	0.0	0.0	16.7	58.3	25.0
	Yitong	YT	0	6	0	7	1	14	0.0	42.9	0.0	50.0	7.1
	Total		0	12	4	32	5	53	0.0	22.6	7.5	60.4	9.4
Yellow	Huating	HT	0	2	0	7	2	11	0.0	18.2	0.0	63.6	18.2
	Luanchuan	LC	4	6	2	27	0	39	10.3	15.4	5.1	69.2	0.0
	Qiaotou	QT	0	4	1	17	0	22	0.0	18.2	4.5	77.3	0.0
	Tantou	TT	7	2	2	16	5	32	21.9	6.3	6.3	50.0	15.6
	Total		11	14	5	67	7	104	10.6	13.5	4.8	64.4	6.7
Huaihe	Beimiaoji	BM	0	0	0	0	12	12	0.0	0.0	0.0	0.0	100.0
	Dapoling	DP	0	6	1	5	9	21	0.0	28.6	4.8	23.8	42.9
	Huangnizhuang	HN	1	0	1	4	4	10	10.0	0.0	10.0	40.0	40.0
	Lixin	LX	0	5	5	4	4	18	0.0	27.8	27.8	22.2	22.2
	Luzhuang	LZ	1	0	0	4	6	11	9.1	0.0	0.0	36.4	54.5
	Peihe	PH	5	0	1	5	7	18	27.8	0.0	5.6	27.8	38.9
	Qilin	QL	2	0	0	1	7	10	20.0	0.0	0.0	10.0	70.0
	Xiagushan	XG	3	3	1	3	9	19	15.8	15.8	5.3	15.8	47.4
	Xinxian	XX	3	3	2	2	14	24	12.5	12.5	8.3	8.3	58.3
	Yangzhuang	YZ	0	5	1	2	2	10	0.0	50.0	10.0	20.0	20.0
	Zhongtang	ZT	2	3	1	4	5	15	13.3	20.0	6.7	26.7	33.3
	Zhuganpu	ZG	4	2	1	2	17	26	15.4	7.7	3.8	7.7	65.4
	Ziluoshan	ZL	3	2	2	8	6	21	14.3	9.5	9.5	38.1	28.6
	Total		24	29	16	44	102	215	11.2	13.5	7.4	20.5	47.4
Yangtze	Anhe	AH	5	3	2	3	1	14	35.7	21.4	14.3	21.4	7.1
	Anren	AR	8	14	3	3	5	33	24.2	42.4	9.1	9.1	15.2
	Baitugang	BT	1	3	1	6	0	11	9.1	27.3	9.1	54.5	0.0
	Biyang	BY	1	1	0	10	0	12	8.3	8.3	0.0	83.3	0.0
	Chengcun	CC	11	3	9	0	0	23	47.8	13.0	39.1	0.0	0.0
	Dutou	DT	6	8	1	8	0	23	26.1	34.8	4.3	34.8	0.0
	Gaotan	GT	4	5	4	6	4	23	17.4	21.7	17.4	26.1	17.4
	Jiahe	JH	6	6	1	0	0	13	46.2	46.2	7.7	0.0	0.0
	Jiajiafang	JJ	2	4	0	4	1	11	18.2	36.4	0.0	36.4	9.1
	Jinping	JP	3	2	6	2	4	17	17.6	11.8	35.3	11.8	23.5
	Jitan	JT	0	2	2	3	4	11	0.0	18.2	18.2	27.3	36.4
	Juwan	JW	4	3	0	8	1	16	25.0	18.8	0.0	50.0	6.3
	Liangshuikou	LK	24	6	6	26	3	65	36.9	9.2	9.2	40.0	4.6
	Liqingdian	LQ	0	6	2	14	7	29	0.0	20.7	6.9	48.3	24.1
	Loudi	LD	7	5	6	2	5	25	28.0	20.0	24.0	8.0	20.0
	Miping	MP	3	3	5	3	5	19	15.8	15.8	26.3	15.8	26.3
	Pingshi	PS	5	3	1	8	5	22	22.7	13.6	4.5	36.4	22.7
	Shahebu	SH	3	3	2	2	0	10	30.0	30.0	20.0	20.0	0.0
	Shanggao	SG	10	2	2	3	2	19	52.6	10.5	10.5	15.8	10.5
	Shijie	SJ	3	4	0	4	2	13	23.1	30.8	0.0	30.8	15.4
	Shimenkan	SM	16	25	2	5	2	50	32.0	50.0	4.0	10.0	4.0
	Shuangfeng	SF	9	8	7	8	1	33	27.3	24.2	21.2	24.2	3.0
	Shuangjiangkou	SK	8	3	12	1	0	24	33.3	12.5	50.0	4.2	0.0
	Sifen	SI	4	2	2	0	2	10	40.0	20.0	20.0	0.0	20.0
	Tangdukou	TD	10	19	1	2	1	33	30.3	57.6	3.0	6.1	3.0
	Tanghe	TH	0	3	1	5	9	18	0.0	16.7	5.6	27.8	50.0
	Tonggu	TG	5	2	0	0	10	17	29.4	11.8	0.0	0.0	58.8
	Tongtang	TO	14	6	5	2	1	28	50.0	21.4	17.9	7.1	3.6
	Wuxigou	WX	4	5	0	7	1	17	23.5	29.4	0.0	41.2	5.9

	Xiawan	XW	6	0	0	2	3	11	54.5	0.0	0.0	18.2	27.3
	Xixia	XI	1	1	3	5	6	16	6.3	6.3	18.8	31.3	37.5
	Xupu	XP	12	14	4	5	1	36	33.3	38.9	11.1	13.9	2.8
	Yanling	YL	18	4	4	7	0	33	54.5	12.1	12.1	21.2	0.0
	Yanta	YA	6	2	1	4	0	13	46.2	15.4	7.7	30.8	0.0
	Yuanken	YK	2	3	1	0	7	13	15.4	23.1	7.7	0.0	53.8
	Yucun	YC	12	0	18	3	1	34	35.3	0.0	52.9	8.8	2.9
	Yuexi	YX	14	4	11	5	3	37	37.8	10.8	29.7	13.5	8.1
	Zhangdou	ZD	4	3	0	5	0	12	33.3	25.0	0.0	41.7	0.0
	Total		251	190	125	181	97	844	29.7	22.5	14.8	21.4	11.5
Southeast	Anxi	AX	1	3	4	6	0	14	7.1	21.4	28.6	42.9	0.0
	Longshan	LS	1	3	16	3	0	23	4.3	13.0	69.6	13.0	0.0
	Tunxi	TX	5	3	1	1	3	13	38.5	23.1	7.7	7.7	23.1
	Xufan	XF	1	3	5	1	0	10	10.0	30.0	50.0	10.0	0.0
	Zhaoan	ZA	1	5	12	8	4	30	3.3	16.7	40.0	26.7	13.3
	Total		9	17	38	19	7	90	10.0	18.9	42.2	21.1	7.8
Pearl	Hezikou	HZ	42	17	7	22	1	89	47.2	19.1	7.9	24.7	1.1
	Huishui	HS	3	3	0	4	0	10	30.0	30.0	0.0	40.0	0.0
	Libo	LB	5	0	0	6	0	11	45.5	0.0	0.0	54.5	0.0
	Xiaogulu	XL	2	24	0	0	4	30	6.7	80.0	0.0	0.0	13.3
	Total		52	44	7	32	5	140	37.1	31.4	5.0	22.9	3.6
Total			347	306	195	375	223	1446	24.0	21.2	13.5	25.9	15.4

Reference

Ball, G. H., and Hall, J.: ISODATA: A Novel Method of Data Analysis and Pattern Classification, Stanford Research Institute, Menlo Park, NTIS No. AD 699616, 1965.

Calinski, T., and Harabasz, J.: A dendrite method for cluster analysis, Communications in Statistics-Theory and Methods, 3, 1-27, https://doi.org/10.1080/03610927408827101, 1974.

Davies, D. L., and Bouldin, D. W.: A cluster separation measure, IEEE Transactions on Pattern Analysis and Machine Intelligence, 1, 224-227, https://doi.org/10.1109/TPAMI.1979.4766909, 1979.

Duda, R. O., and Hart, P. E.: Pattern Classification and Scene Analysis, John Wiley & Sons, New York, ISBN 0471223611, 1973.

Dunn, J. C.: Well-separated clusters and optimal fuzzy partitions, Journal of Cybernetics, 4, 95-104, https://doi.org/10.1080/01969727408546059, 1974.

Friedman, H. P., and Rubin, J.: On some invariant criteria for grouping data, Journal of the American Statistical Association, 62, 1159-1178, https://doi.org/10.1080/01621459.1967.10500923, 1967.

Friedman, H. X., and Rubin, J.: On some invariant criteria for grouping data, Journal of the American Statistical Association, 62, 1159-1178, https://doi.org/10.2307/2283767, 1967.

Halkidi, M., and Vazirgiannis, M.: Clustering validity assessment: Finding the optimal partitioning of a data set, in: Proceedings 2001 IEEE International Conference on Data Mining, San Jose CA, USA, 29 November-02 December 2001, 187-194, https://doi.org/10.1109/ICDM.2001.989517, 2001.

Halkidi, M., Vazirgiannis, M., and Batistakis, I.: Quality scheme assessment in the clustering process, in: Principles of Data Mining and Knowledge Discovery: 4th European Conference, PKDD 2000 Lyon, France, 13-16 September 2000, 265-276, 2000.

Hartigan, J. A.: Clustering Algorithms, John Wiley & Sons, New York, ISBN 047135645X1975, 1975.

Hubert, L. J., and Levin, J. R.: A general statistical framework for assessing categorical clustering in free recall, Psychological Bulletin, 83, 1072-1080, https://doi.org/10.1037/0033-2909.83.6.1072, 1976.

Krzanowski, W., and Lai, Y.: A criterion for determining the number of groups in a data set using sum-of-squares clustering, Biometrics, 44, 23-34, https://doi.org/10.2307/2531893, 1988.

Marriott, F. H. C.: Practical problems in a method of cluster analysis, Biometrics, 27, 501-514, https://doi.org/10.2307/2528592, 1971.

McClain, J. O., and Rao, V. R.: Clustisz: A program to test for the quality of clustering of a set of objects, Journal of Marketing Research, 12, 456-460, https://doi.org/10.2307/3151097, 1975.

Milligan, G. W., and Cooper, M. C.: An examination of procedures for determining the number of clusters in a data set, Psychometrika, 50, 159-179, https://doi.org/10.1007/BF02294245, 1985.

Milligan, G. W.: A Monte Carlo study of thirty internal criterion measures for cluster analysis, Psychometrika, 46, 187-199, https://doi.org/10.1007/BF02293899, 1981.

Milligan, G. W.: An examination of the effect of six types of error perturbation on fifteen clustering algorithms, Psychometrika, 45, 325-342, https://doi.org/10.1007/BF02293907, 1980.

Ratkowsky, D. A., and Lance, G. N.: Criterion for determining the number of groups in a classification, Australian Computer Journal, 10, 115-117, 1978.

Rousseeuw, P. J.: Silhouettes: a graphical aid to the interpretation and validation of cluster analysis, Journal of Computational and Applied Mathematics, 20, 53-65, https://doi.org/10.1016/0377-0427(87)90125-7, 1987.

Sarle, W. S.: SAS Technical Report A-108, Cubic Clustering Criterion, SAS Institute Inc, Cary, NC, 1983.

Scott, A. J., and Symons, M. J.: Clustering methods based on likelihood ratio criteria, Biometrics, 27, 387-397, https://doi.org/10.2307/2529003, 1971.