Physical controls of regional distribution patterns of precipitation and flow duration curves in the middle and lower reaches of the Yangtze River

Yu Zhou[1] and Yu-Jia Zhang[1]

[1]Zhejiang University

May 5, 2023

Abstract

The flow duration curve (FDC) is the cumulative distribution function, which represents the relationship between the frequency and magnitude of streamflow,and the precipitation duration curves (PDC) follows the same principle. Nowadays, the correlation between the shape of PDC and FDC curves, their respective physical control factors, and their fitting conditions in unmeasured catchments across China have not been fully understood. In this paper, daily precipitation from 698 weather stations across China and streamflow from more than 200 hydrological stations in the middle and lower Yangtze River basin were chosen to analyze the relationship, similarity, regional pattern and response mechanism of fitting parameters between PDC and FDCs. Framework was proposed for modeling FDC, decomposing the Streamflow time series into fast flow and slow flow time series and attributing the shapes of PDC and FDCs to catchment meteorological and geographical characteristics and physical processes. Results indicate that the parameters of PDC and certain FDCs (TFDC, FFDC, SFDC) share similar spatial patterns but the value of parameters and shape of curves varies for the different duration and interactions of the processes. The climate and catchment characteristics such as extreme properties of precipitation, base flow index (*BFI*), *Πμαξ*απ* and concentration ratio index based on monthly precipitation (*CIM*) will influence the shape of normalized PDC and FDCs, which provides a way to predict unmeasured catchments for PDC and FDCs, diagnose catchment rainfall-runoff responses, including similarity and differences between catchments, and can be applied to more future research about processes based on physical controls.

1. INTRODUCTION

Flow duration curves (FDC) is a statistical expression of the flow changes observed during the period of record (long-term flow duration curves or annual flow duration curves) based on the daily runoff series, which explains the relationship between the flow size and the probability of occurrence (Ghotbi et al., 2020; Leong and Yokoo, 2021). The probability value of a specified flow equal to or exceeding a certain flow size can be read from FDC, so it is essentially the cumulative distribution function of daily flow (Yokoo and Sivapalan, 2011). There are two options about the data recording time of FDC: the total duration method (Cheng et al., 2012) and the multi-year average method (Cheng et al., 2012; Karst et al., 2019; Liang, 2019). Cheng et al. used the three-parameter mixed gamma distribution to characterize the shape of the normalized FDC using the total data record, and observed the change of FDC year by year in eight representative catchments. The total duration method more closely represents the rainfall and runoff of a specific basin, while the fitting curve calculated by the multi-year average method can fully capture the changes of FDC between years (Cheng et al., 2012; Costa and Fernandes, 2021; Liang, 2019). FDC can be used to diagnose the rainfall-runoff response of a watershed to help develop or validate rainfall-runoff models through the transition from precipitation change to runoff (Wang et al., 2023). It can be also used to analyze the similarity and difference between watersheds, establish a model using precipitation and watershed characteristics as input, and predict the hydrological elements of ungauged sub-watersheds from a limited number of stations observed in the past (Burgan and Aksoy, 2022a; Wolff and Duarte, 2021).

In the past, the physical control of FDC involved graphics (non-parametric method), statistics (parametric method) and process-based methods (Ghotbi et al., 2020; Ghotbi et al., 2020; Karst et al., 2019; Müller and Thompson, 2016). The graphic method is based on exploring the relationship between the shape characteristics of FDC and the climate and geomorphologic characteristics of the catchment, such as the steepness or quantile of the curve, to estimate the shape of the FDC which represents the flow condition of unmeasured catchments (MOHAMOUD, 2008). The statistical method aims to fit the empirical FDC through appropriate distribution functions (such as gamma distribution, lognormal distribution, generalized Pareto distribution, kappa distribution, etc.) (Almeida et al., 2021; Burgan and Aksoy, 2022b; Cheng et al., 2012; Ghotbi et al., 2020; Yire Shin, 2022), and then find the quantitative and qualitative relationship between the estimated parameters of the fitting function and the characteristics of the local climate and geographical environment of the measured catchment. Villalobos and Neelin explained why the gamma distribution can well fit the daily precipitation distribution. The control of climate and watershed geographical environment characteristics on FDC always tends to be empirical (Ridolfi et al., 2020). Based on these empirical properties and the diversity of climate and geographical impact factors in different watersheds, the main hydrological processes are not explicitly included. The main hydrological processes are complex and specific in different watersheds and climatic regions, so the follow-up research is mostly the process-based method. The essence of the process-based method is to combine graphic with statistical methods, take into account the dominant process of time flow change, take rainfall series as random input, take seasonal and other meteorological factors as random influence, establish a deterministic model of runoff involving precipitation, and derive the explicit statistical distribution form of FDC (Chouaib et al., 2018; Leong and Yokoo, 2021).

Among these three methods, the parametric method shows that the geographical characteristics of the catchment will affect the shape of FDC. For example, topographic features, including the area, average elevation and gradient of the catchment area (Luan et al., 2021; Yang et al., 2023), have a great impact on the shape of FDC. Stephen Oppong Kwakye and András confirmed the control effect of rainfall events and dry conditions on the FDC of most rivers in West Africa (Kwakye and Bárdossy, 2022). Cheng et al. found that in the control of 197 watersheds in the United States, daily precipitation, probability of non-rainy days and Aridity index (AI) has a great impact on the parameter a estimated by gamma distribution. Although there may be significant differences in the key control factors of FDCs at different geographical locations and times, the shape of FDCs represents the comprehensive impact of rainfall, topography, soil, geology and other climatic and geomorphologic characteristics.

The main problem of the three methods is that the process of precipitation input and runoff generation involves different time scales (Huang et al., 2020). For example, the runoff generation process usually includes fast land flow and groundwater flow. The fast flow is the hydrological response to the change of precipitation and the result of land flow generated by the mechanism of excess infiltration, which is related to many factors such as rainfall intensity, soil infiltration capacity and early soil moisture content, etc. While the slow flow is related to climate seasonality and is controlled by soil permeability and terrain slope. Therefore, when establishing hydrological models and performing numerical simulation, the total flow was always divided into two parts: fast flow and slow flow, so as to have a deeper understanding of the shape characteristics and process control of FDC. Based on the work of Yokoo and Sivapalan (2011) and Cheng et al. Ghotbi et al. (2020) proposed a new stochastic framework to construct total flow, involving three parts of partition (total, fast and slow flow), fitted their distributions and then combined them together as an FDC model of total flow. In this way, the FDC is usually divided into total flow duration curve (TFDC), fast flow duration curve (FFDC) and slow flow duration curve (SFDC). However, the fast flow is usually highly intermittent, while the slow flow is usually more continuous, which poses a challenge to explain the statistical dependence between slow flow and fast flow. In addition, in previous work, Ghotbi et al. (2020) did not further implement it to explore the climate and landscape controls of FDC. At present, most research about FDC is distributed in many countries such as the United States, Africa and Italy (Chouaib et al., 2019; Cislaghi et al., 2020; Kwakye and Bárdossy, 2022; Over T M, 2018; SMAKHTIN et al., 1997; Ye et al., 2012). Due to China's vast area, complex terrain, and many unmeasured catchment areas, it is very meaningful to study the FDC of unmeasured catchments across China, and provide a good theoretical and practical basis for its research.

In this paper, based on the continuous daily precipitation data (56yr) of 698 weather stations in China and streamflow data (30yr) of 244 gauging stations in watersheds in the middle and lower Yangtze River basin, we carried out the gamma distribution fitting and parameter estimation of PDC and FDCs, and explored the internal and external interfering factors of each fitting parameter. Firstly, data sources were described for the analyses and a brief overview of methodology adopted were gave. Secondly, the fitting images of PDC and FDCs were shown and the values of R^2 and Nse were calculated to evaluate the fitting effect. Then the spatial distribution pattern of fitting parameters which controls the shape of PDC and FDCs were presented. Finally we analyzed the estimated parameters with their correlation with regional meteorological physical characteristics, took the middle and lower Yangtze River basin as an example with base flow segmentation, and studied the relationship, similarity, regional patterns and response mechanism of parameters between PDC and FDCs.

2. DATA AND METHODS

2.1 Data

In this work, Daily precipitation and temperature data were downloaded from 698 evenly distributed weather stations nationwide from the China Meteorological Data Network from 1961 to 2016 (*https://data.cma.cn/*), which covers a wide range of climate, ecological regions and landscapes throughout China. The temperature data was used to estimate the potential evaporation. Then abnormal data was excluded, and the research parameter changes and relationships in the middle and lower reaches of the Yangtze River were focused on. The 30 meter digital elevation model (DEM) was downloaded from the geospatial data cloud (*http://www.gscloud.cn/*) by using ArcGIS, and the corresponding watershed area of the hydrological station was extracted. According to the Hydrological Annual Report of the People's Republic of China, daily runoff data was collected from 224 stations with records from 1970 to 1990 and from 2007 to 2016 for at least 20 years. Daily precipitation and temperature data were chosen from 267 weather stations in and near Yangtze river basin (YZRB). The observed precipitation and estimated potential evaporation were interpolated into the whole YZRB using the Thiessen polygon method (Meena et al., 2013). Then the interpolated precipitation and potential evaporation of the basin area corresponding to 224 hydrological stations were averaged to obtain the daily precipitation and potential evaporation data of each hydrological station.

2.2 Definition of dry spell and wet spell

According to the research conducted by Bichet et al. (Bichet and Diedhiou, 2018) when there are two or more consecutive dry days with precipitation less than 0.1 mm or 1 mm, the multi-year average days of the maximum continuous day without precipitation was defined as a dry spell. Similarly, when there are two or more consecutive wet days with precipitation greater than 0.1 mm or 1 mm, and the multi-year days of average maximum continuous precipitation day is defined as a wet spell. In addition, Pendergrass pointed out that different thresholds (0.1mm or 1mm) might lead to different conclusions during the analysis of changes in precipitation characteristics (Pendergrass, 2018). In this article, 1 mm is selected as the threshold for calculation.

2.3 Gamma distribution fitting of hydrological sequences with zero values

Generally speaking, if PDC is represented by an overly complex distribution with multiple parameters, a more accurate curve fitting may be obtained. However, some correlation between different parameters may exist, leading to uncertainty in parameter estimation, which may confuse the physical control on statistical parameters. Therefore, in order to achieve the goal of this paper, a simple statistical distribution (the gamma distribution) was selected. Considering the need for parameter simplicity and the need to link these parameters with climate attributes, a three-parameters continuous probability distribution is adopted, which is determined by the probability of occurrence of (zero value) , shape parameters a and scale parameters β . PDC and FDC are complementary cumulative distribution functions (CCDF) of daily precipitation and daily runoff respectively, which must adapt to the presence of flow (zero value condition), especially in arid areas. Therefore, in this study, we used the following gamma distribution to represent FDC, as shown in Eq. .

Where is the probability of occurrence of zero value (precipitation, flow, etc.); is the probability density function of gamma distribution, and is defined as Eq.

in which α $\alpha\nu\delta$ β are the shape and scale parameters respectively. The probability of exceedance distribution P can be calculated by Eq. , where G^{-1} is the inverse function of the CCDF.

According to the above formula, controls the zero part of the duration curve, while α and β control the shape of the non-zero part of the duration curve.

The scale parameter β has great impacts on the vertical offset of PDC/FDC. The larger the observed average flow rate, the higher the valueβ . In addition, the shape parameter α essentially controls the slope of FDC. The smaller the value α , the steeper the slope of FDC. The parameter was directly estimated from the observation results, which is the fraction of the number of days with zero precipitation (flow) in the data record, that is, the number of days with zero precipitation (flow) divided by the total number of days in the precipitation (flow) record. The parameters α and β of the function were estimated using the moment method (Saulo et al., 2018) through the relationship between the mean (μ) and variance (v) of the function distribution (Eqs. and).

and are estimated from time series with q >0 (non-zero part of FDC). In addition, R-squared (R^2) and Nash Sutcliffe efficiency coefficient (Nse) (Nash and Sutcliffe, 1970) were selected to evaluate the performance of gamma distribution in providing good fitting for non-zero segments of different duration curves in each catchment becauseR^2 (Eq.) only measures the degree of linear correlation, while Nse (Eq.) represents the matching degree between observations and estimates, and explains the bias at the same time.

where and are the predicted and the observed value, respectively;n represents the total number of days included in those years when representing the long-term time curve, and represents the number of days included in one year when representing the one-year time curve, respectively. and are the average values of observed and predicted values. When the value of R^2 and Nse are close to 1, it represents that the gamma distribution fits well.

2.4 Digital filtering method for separating base flow

The digital filtering method is based on signal analysis and processing technology, whose principle is to decompose daily flow data into high- and low-frequency signals. It is believed that surface flow is a high-frequency signal while the base flow is a low-frequency signal, thus fast flow and base flow can be separated accordingly (Nathan and McMahon, 1990). In this paper, Lyne Hollick filtering method (Eqs. and) was adopted to separate.

where is the surface flow (fast flow) at time t , while and is the total flow and base flow at the current time t respectively. is the filtering parameter, with the value of 0.925 (Chapman, 1991).

By applying the baseflow separation methods to daily streamflow, daily streamflow are schematically partitioned into separate fast flow and slow flow time series, as shown in **Figure 1** (a). Processes controlling fast flows are surface runoff generation and routing. The variability of fast flows is governed by stochastic characteristics of the sequences of storm events experienced by the catchment and the properties of surface soils and topography. Processes controlling slow flow include subsurface flow and groundwater discharge. The variability of slow flows strongly reflects climate seasonality and the underlying geology of the aquifer system. In this paper, the total flow, fast flow and slow flow series were normalized with the value of average daily total flow, average daily fast flow and average daily slow flow respectively. In addition, a normalized time series (i.e. daily flow divided by long-term average daily flow) was used to construct the empirical duration curve: TFDC, FFDC and SFDC (Yokoo & Sivapalan, 2011), as demonstrated in **Figure 1** (b).

Hosted file

`image25.emf` available at https://authorea.com/users/614962/articles/641593-physical-controls-of-regional-distribution-patterns-of-precipitation-and-flow-duration-curves-in-the-middle-and-lower-reaches-of-the-yangtze-river

Figure 1 The schematic illustration of the proposed framework for modeling flow duration curve (FDC): (a) Streamflow time series is decomposed into fast flow time series and slow flow time series; and (b) FDC is computed as the sum of fast flow and slow flow considering the dependence between them.

3. RESULTS

3.1 Performance of the mixed gamma Distribution

Normalized sequences were used to construct PDC, TFDC, FFDC, and SFDC for 698 weather stations and 224 hydrological stations. Under interannual variations, annual precipitation records for 56 years and annual flow records for at least 20 years were constructed, respectively. According to respective normalized sequences, gamma distribution was performed on each duration curve and moment method were used to determine model parameters α and β . The estimated parameter , α , β were used to predict the PDC and FDCs via formula (3). In this paper, the Nash-Sutcliffe efficiency coefficient (Nse) and the Goodness of fit (R^2) of PDC and FDCs fitted by gamma distribution were used to evaluate the fitting effect and accuracy.

For each duration curve, the aggregated statistical data of $Nse$and R^2 obtained from the mixed gamma distribution fitting of normalized duration curves for 698 weather stations and 224 hydrological stations are shown in **Figure 3** . The upper and lower whiskers represent the maximum and minimum values and the upper and lower lines of the box represent the third and first quartiles. It can be seen from the **Figure 3** that although visual differences of selected sites based on typical climate conditions are found in **Figure 2** , the mixed gamma distribution still fits the empirical duration curves: PDC, TFDC, FFDC and SFDC well.

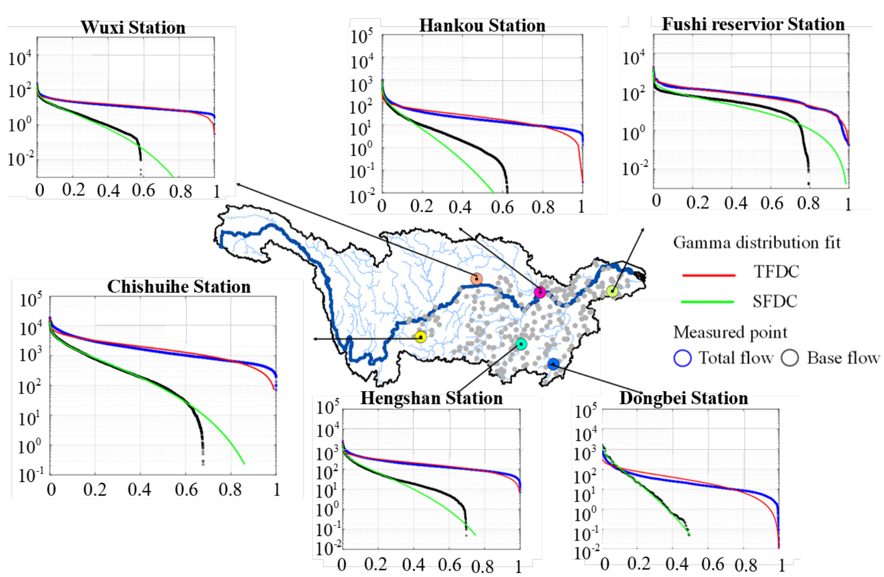


Figure 2 Normalized empirical FDCs (point) and gamma distribution fits (solid lines) for 6 randomized stations: TFDC (red lines, blue points) and SFDC (green lines, black points).

Hosted file

`image28.emf` available at https://authorea.com/users/614962/articles/641593-physical-controls-of-regional-distribution-patterns-of-precipitation-and-flow-duration-curves-in-the-middle-and-lower-reaches-of-the-yangtze-river

Figure 3 Values ofR^2 and *Nse* of the mixed gamma distribution fitting to normalized duration curves. The upper and lower whiskers represent the maximum and minimum values and the upper and lower lines of the box represent the third and first quartiles.

Among 698 weather stations and 224 hydrological stations, the 75% of the estimated values of*Nse* and R^2 exceed 0.81 and the minimum values are greater than 0.6. This reflects that the mixed gamma distribution can perfectly capture the shape of PDC and FDCs.

From different hydrological stations' distribution of *Nse* andR^2 in FDC, it can be seen that there is a difference. The values of R^2 and *Nse* of some stations are relative low, indicating that the fitting effect of FDC is not very good at certain hydrological stations or in certain years. On the other hand, the fitting performance of PDCs and FFDCs is superior to that of TFDCs and SFDCs. It may be due to the fact that the gamma fitted data should meet the requirements of independently identically distribution, which presents the characteristic of the random-like properties. Precipitation and surface runoff processes include more complex stochastic processes, which are more able to reflect this randomness.

3.2 Regional patterns in spatial variation of PDC

Hosted file

`image29.emf` available at https://authorea.com/users/614962/articles/641593-physical-controls-of-regional-distribution-patterns-of-precipitation-and-flow-duration-curves-in-the-middle-and-lower-reaches-of-the-yangtze-river

Figure 4 Spatial distribution of precipitation estimation parameters α , $\beta$for 698 weather stations in China

As shown in **Figure 4** . the gamma distribution is used to well fit the estimated parameters α and β of PDC for 698 weather stations across China. The values of α , except for the stations in the southwest and southeast inland regions of China distributed above 0.6, is basically distributed between 0.2 and 0.6. From **Figure 4** (b), it can be found an increasing trend ofβ from northwest to southeast of China, and β is mainly distributed between 0 and 10 in northwest China, and is mainly distributed between 10-20 at stations in Northeast China, Central China, and Yunnan Province, and is distributed between 20-30 in most areas in the southeast, with a small number of coastal cities in the southeast distributed between 30 and 70. From the figure, it is discovered thatβ shares a similar spatial distribution pattern with average annual precipitation in China.

3.3 Regional patterns in spatial variation of FDCs

The estimation parameter α and β of FDCs at 224 hydrological stations were fitted and calculated in the middle and lower reaches of the Yangtze River Basin,then α and β of PDC of these stations in corresponding watersheds were calculated too using the Tyson polygon method. As depicted in **Figure 5** , α andβ of different flow duration curves present different regional distribution patterns.

According to gamma distribution estimation parameters, in stochastic process, shape parameter α represents the number of events, while the scale parameter β indicates the occurrence rate of the event, i.e β =,For a fixed time rate, if we expect more time to occur, i.e the larger the α , the longer the waiting time T will be; For a fixed number of events , when the event rate is high, it means that the smaller β , the shorter the waiting time T will be.

From **Figure 5** , it can be seen that the value α of FFDCs is basically smaller than that of other duration curves. Among the four duration curves, the value α of SFDC is relatively larger, due to the shorter duration time of fast flow and longer duration time of slow flow. The steeper the slope of the area, the shorter the diffusion time of the duration curve, that is the smaller the value α will be. In contrast, the spatial distribution of the values α of PDCs and FFDCs and those of TFDCs and SFDCs have a similar regional pattern. However, the value α of PDCs is generally greater than FFDCs, and the value α of SFDCs is generally greater than TFDCs. What's more, the range of value α of PDC and FFDCs is much narrower than that of TFDCs and SFDCs, which is consistent with Cheng's research. Similarly, among the four duration curves, the value β of FFDCs is relatively smaller, due to the higher occurrence rate of fast flow, considering the shorter duration of fast flow and the longer duration of slow flow.

Overall, the regional patterns of α and β in 224 hydrological stations in the **Figure 5** indicate that they are influenced by climate differences, but their differences within the same spatial patterns related to climate zones are also significant. This may indicate that other secondary climate and physiological characteristics may also be the reasons for the differences in FDC shape. Possible factors could include climate seasonality, groundwater contribution, vegetation, slope, and watershed shape. The following analysis aims to determine which of these controls may dominate.

Hosted file

`image34.emf` available at https://authorea.com/users/614962/articles/641593-physical-controls-of-regional-distribution-patterns-of-precipitation-and-flow-duration-curves-in-the-middle-and-lower-reaches-of-the-yangtze-river

Figure 5 Regional patterns of variation of parameter α for the different duration curves(a) PDC; (b) TFDC (C) FFDC; (d)SFDC.

4. DISCUSSION

4.1 Τηε ςορρελατιον βετωεεν β ανδ εξτρεμε ςηαραςτεριστιςς οφ πρεςιπιτατιον

In order to further explore the parameters α and β of PDC on a national scale, we analyzed the Spearman's rank correlation coefficient (Spearman, 2010) between them and the physical control factors to obtain their possible influencing factors. As shown in**Figure 6** , β is inversely proportional to the Aridity index (AI), and has a high correlation with P and precipitation percentile, with a correlation of 0.9816 with 99 wet day percentile precipitation. It further confirms the conclusion of Geng, S to some extent that catchments with higher precipitation tend to have larger values of β which are estimated by gamma distribution (Geng et al., 1986). By comparing **Figure 6** (c) (d) (e) and (f), it can be seen that the larger the values of β , the more dispersed the probability distribution of precipitation at different levels is, with the greater likelihood of extreme precipitation occurrence.

Schär C et al. pointed out that the precipitation percentile (or quantile) was used to evaluate the trend and prediction of heavy precipitation events, and analyzed the temporal and spatial distribution and change characteristics of extreme precipitation threshold (Schär et al., 2016). In evaluating extreme events such as rainstorm, the use of events such as drizzle (precipitation less than 0.1 mm or 1 mm) seems illogical, so the wet day precipitation threshold has a better performance on the extreme precipitation threshold of heavy precipitation events.

Hosted file

`image35.emf` available at https://authorea.com/users/614962/articles/641593-physical-controls-of-regional-distribution-patterns-of-precipitation-and-flow-duration-curves-in-the-middle-and-lower-reaches-of-the-yangtze-river

Figure 6 Estimated Parameters β 's relationship with various influencing factors

It can be seen from **Figure 6** (a) (d) (f), that when comparing the correlation between β and precipitation, 95th percentile precipitation, and 99th percentile precipitation, the scale parameterβ is more closely related to high precipitation, with the greater correlation with extreme precipitation.

4.2 Similarity among different duration curves

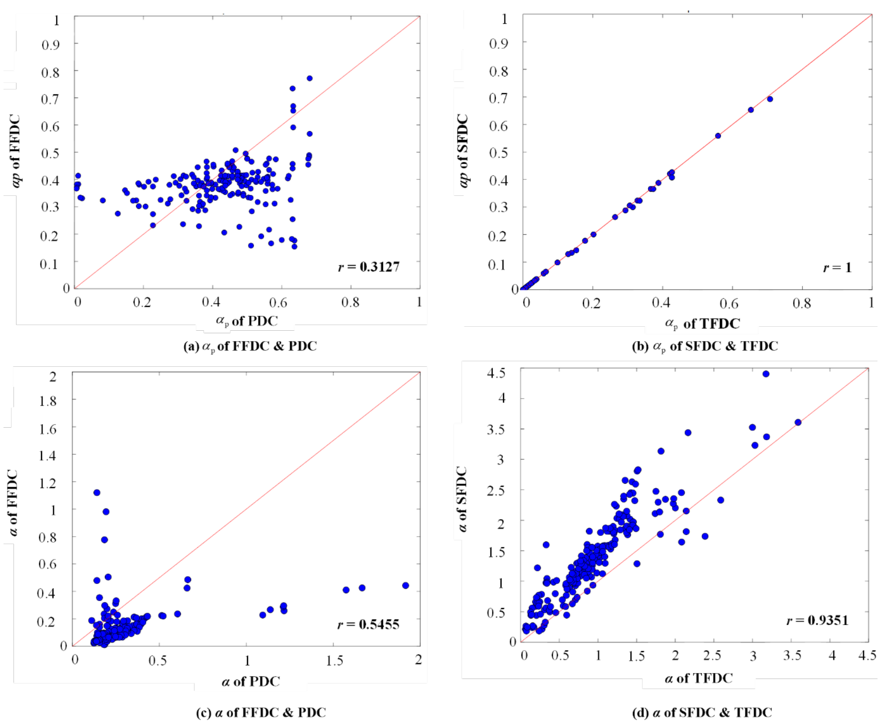


Figure 7 Cross-correlation of parameters p_0 and $a$of the mixed gamma distribution between PDC and FFDC ((a) and (c)), and between TFDC and SFDC ((b) and (d)).

The values of parameters$a_{\text{ p}}$ and a of PDC and FFDC have similar spatial patterns, and those of TFDC and SFDC also exhibit similar spatial patterns (**Figure 5** and **Figure 6**). As shown in**Figure 7** , compared with TFDC and SFDC, the parameter$a_{\text{ p}}$ of PDC and FFDC is correlated, but the deviation is much greater. The proportion of zero flow days ($a_{\text{ p}}$) of FFDCs can be converted from the proportion of PDC on the daily time scale, but it may be a non-linear relationship (Yokoo and Sivapalan, 2011). The parameter$a_{\text{ p}}$ of FFDC is related to the total days of non-precipitation, that is, the p_0 parameter of PDC. However, fast flow only occurs when precipitation meets the initial loss, field capacity of soil, and/or exceeds the infiltration capacity of the surface soil layer. In addition, after fast flow is generated., watershed characteristics such as slope, shape, and water system will have additional effects. Therefore, there is a correlation between PDC and FFDCs' parameters, but there also exist a certain degree of dispersion, which is a challenge for us to infer the parameters of FFDC based on existing precipitation data. The spatial distribution of the value a of PDC and FFDC and that of TFDC and SFDC have a similar regional pattern (**Figure 5**). It can also be seen from**Figure 7** that there is a strong correlation between PDCs and FFDCs, as well as between TFDC and SFDC. However, the similarity in shape between TFDC and SFDC is more related than that between PDC and FFDC, with the former having the Spearman's rank correlation coefficient of 0.9351 while the latter having the correlation coefficient of 0.5455. The discrete relationship of parameters a between PDC and FFDC indicates that other factors may also have an impact on the correlation between PDC and FFDC, including terrain effects. The response of fast flow to precipitation is influenced by the terrain characteristics of the watershed, including vegetation coverage, gradient, and soil characteristics. Because little rain will not generate fast flow, the similarity between PDC and FFDCs' lower tail is weak,FFDC is steeper, and the value a of FFDC is generally lower than that of PDC.

For the similarity between TFDC and SFDC, the parameter α of SFDC is generally greater than that of TFDC, because higher flow is separated and the duration of the slow flow is longer. Therefore, the SFDC is flatter and its α is larger. The shape difference between TFDC and SFDC also appears at the bottom of the duration curves. It can be discovered from **Figure 7** (c) and (d) that the linear relationship is good distributed within smaller values, reflecting both similarities decrease as αincrease. The conclusions of spatial pattern research on four duration curves' parameters may not be directly applied to practical life, but it is conducive to more future work about processes based on physical influencing factors.

4.3 Correlating parameters with catchment physical characteristics

Regression was conducted with the shape parameters α of TFDC, FFDC, and SFDC and some climate and geographical features. As shown in**Figure 8** (a) and (b), it is clear that the parameters $\alpha$of TFDC and SFDC is closely related to the base flow index (BFI) between watersheds, where BFI is a common indicator of the response features of runoff in watersheds, reflecting the comprehensive impact of local geographical and climate characteristics. In addition, it can be found that the shape changes of TFDC and SFDC at 224 hydrological stations monotonically increase with the increase ofBFI , which means that the spatial changes of TFDCs and SFDCs are controlled by climate and watershed characteristics that affect the division of runoff into fast and slow flow. Watersheds with largerBFI should have a larger proportion of slow flow, so the curves of TFDC or SFDC tend to be flatter. Therefore, BFI can be used to estimate the value of α .

Hosted file

`image37.emf` available at https://authorea.com/users/614962/articles/641593-physical-controls-of-regional-distribution-patterns-of-precipitation-and-flow-duration-curves-in-the-middle-and-lower-reaches-of-the-yangtze-river

Figure 8 Correlation of parameter α of TFDC, SFDC and FFDC of 224 catchments with catchment physical characteristics BFI ,$P_{\text{max}*}$ α_{p}.

As shown in **Figure 8** (c), α of FFDC is closely related to the multiplication of maximum daily precipitationP_{max} and α_{π} . Due to precipitation being a prerequisite for fast flow andP_{max} basically representing the maximum intensity of daily precipitation. For specific catchments, the maximum fast flow usually occurs when the maximum daily precipitation events happen. Moreover, the probability of exceedance of the fast flow depends on the occurrence of precipitation, as a supplementary condition. Therefore,$\Pi_{\mu\alpha\xi^*}\alpha_{\pi}$ is selected as the climatic influencing factors for the shape of FFDC with the Spearman's rank correlation coefficient of -0.4139. the shape parameters $\alpha$of FFDC tends to decrease as$\Pi_{\mu\alpha\xi^*}\alpha_{\pi}$ increases. Based on this, the results of $\Pi_{\mu\alpha\xi^*}\alpha_{\pi}$ can be used to predict the spatial variation of FFDC.

Hosted file

`image38.emf` available at https://authorea.com/users/614962/articles/641593-physical-controls-of-regional-distribution-patterns-of-precipitation-and-flow-duration-curves-in-the-middle-and-lower-reaches-of-the-yangtze-river

Figure 9 Correlation of parameter α of PDC and FFDC of 224 catchments with catchment physical characteristics CIM .

There is a relationship between parameter α of PDC and FFDC of 224 catchments and catchment physical characteristics CIM (**Figure 9**). From **Figure 9** (b), it can be seen that there is a significant correlation (negative correlation) between the seasonal index - Concentration ratio index based on monthly precipitation (CIM) and α of FFDC, but this correlation is not as tight as $\Pi_{\mu\alpha\xi^*}\alpha_{\pi}$. In order to analyze the mechanism by which CIM affects the shape of the PDC curves, we conducted further research on the correlation between values $\alpha$of PDC and CIM based on the previous research conducted in Section 4.1 (α of PDC is not related to the factors mentioned that affect β), which indicates that α of PDC has a tighter correlation with CIM (**Figure 9** (a)). Therefore, based on the previous analysis of the impact of precipitation on fast flow, CIM has impacts on the curve of PDC, which in turn affects the curve shape of FFDC.

For total flow and fast flow, the base flow index (*BFI*) can be considered as the main influencing factor (but not the only) on the shape of normalized FDCs (TFDC and SFDC), $\Pi_{\mu a \xi^* a_\pi}$ influences the shape of FFDC curves. *CIM* indirectly affects the change of fast flow through influencing precipitation. However, because the precipitation process is relatively complex, CIM has more influence on the shape of PDC than on FFDC.

5. CONCLUSIONS

The flow duration curve (FDC) is a hydrologically representation of the statistical distribution of daily streamflow. The complexity of processes contributing to the FDC introduces challenges for the direct exploration of physical controls on FDC. (Pumo et al., 2013; Yokoo and Sivapalan, 2011). In this paper, we used 56-year continuous daily precipitation from 698 weather stations across China and streamflow from more than 200 hydrological stations in the middle and lower Yangtze River basin, proposed a new framework to analyze the physical controls of PDC and FDCs applicable to the catchments in the middle and lower reaches of the Yangtze River, which includes dividing the total flow into fast and slow parts, constructing the FDCs by combining FDCs of fast and slow flow (i.e. FFDC and SFDC, respectively) and performing curves fitting and parameter estimation. We quantified the dependency structure between precipitation and flow to measure the correlation and interactions between different duration curves, which can be controlled by climatic and catchment characteristics.

1. The shape parameters a of PDC are related to CIM while the scale parameters β of PDC has a high degree of correlation with precipitation and precipitation percentiles (The correlation betweenβ and 99th percentile wet day precipitation reaches 0.9816), that is, spatial changes in scale parameters are significantly affected by precipitation patterns, especially by extreme precipitation.

2. The values of parameters p_0 and a of PDC and FFDC have similar spatial patterns, and those of TFDC and SFDC also exhibit similar spatial patterns. but there exists a certain degree of dispersion between PDC and FFDC. The Spearman's rank correlation coefficient between a of TFDC and SFDC is 0.9351 while it is 0.5455 between a of PDC and FFDC because many factors may also have an impact on the correlation between PDC and FFDC, including terrain effects,vegetation coverage, gradient, and soil characteristics.

3. The similarity between PDC and FFDCs' lower tail is weak,FFDC is steeper, and the value a of FFDC is generally lower than that of PDC. Similarly, for the similarity between TFDC and SFDC, the parameter$a$ of SFDC is generally greater than those of TFDC, because higher flow is separated and the duration of the slow flow is longer. Therefore, the SFDC is flatter and its a is larger.

4. For total flow and fast flow, the base flow index (*BFI*) can be considered as the main influencing factor on the shape of normalized FDCs (TFDC and SFDC),$\Pi \mu a \xi^* a \pi$ influences the shape of FFDC curves. CIM indirectly affects the change of fast flow through influencing precipitation. However, because the precipitation process is relatively complex, the impact of CIM on the shape of PDC is greater than that on FFDC.

DATA AVAILABILITY

DEM data was downloaded from Geospatial Data Cloud at http://www.gscloud.cn/. Climatological data used in this study was obtained from China Meteorological Data Network, which can be accessed at http://data.cma.cn/. Discharge data comes from Annual Hydrological Report of the People's Republic of China issued by Yangtze River Water Resources Commission.

REFERENCES

References:

Almeida, M., Pombo, S., Rebelo, R. and Coelho, P., 2021. The probability distribution of daily streamflow in perennial rivers of Angola. Journal of Hydrology, 603: 126869.

Bichet, A. and Diedhiou, A., 2018. West African Sahel has become wetter during the last 30 years, but dry spells are shorter and more frequent. Climate Research, 75(2): 155-162.

Burgan, H.I. and Aksoy, H., 2022a. Daily flow duration curve model for ungauged intermittent subbasins of gauged rivers. Journal of Hydrology, 604: 127249.

Burgan, H.I. and Aksoy, H., 2022b. Daily flow duration curve model for ungauged intermittent subbasins of gauged rivers. Journal of Hydrology, 604: 127249.

Chapman, T.G., 1991. Comment on "Evaluation of automated techniques for base flow and recession analyses" by R. J. Nathan and T. A. McMahon. Water Resources Research, 27(7): 1387-1786.

Cheng, L. et al., 2012. Exploring the physical controls of regional patterns of flow duration curves – Part 1: Insights from statistical analyses. Hydrology and Earth System Sciences, 16(11): 4435-4446.

Chouaib, W., Alila, Y. and Caldwell, P.V., 2019. On the use of mean monthly runoff to predict the flow-duration curve in ungauged catchments. Hydrological sciences journal, 64(13): 1573-1587.

Cislaghi, A., Masseroni, D., Massari, C., Camici, S. and Brocca, L., 2020. Combining a rainfall-runoff model and a regionalization approach for flood and water resource assessment in the western Po Valley, Italy. Hydrological sciences journal, 65(3): 348-370.

Costa, V. and Fernandes, W., 2021. Regional Modeling of Long-Term and Annual Flow Duration Curves: Reliability for Information Transfer with Evolutionary Polynomial Regression. Journal of Hydrologic Engineering, 26(2): 04020067–.

Geng, S., de Vries, F.W.T.P. and Supit, I., 1986. A simple method for generating daily rainfall data. AGRICULTURAL AND FOREST METEOROLOGY, 36(4): 363–376.

Ghotbi, S., Wang, D., Singh, A., Blöschl, G. and Sivapalan, M., 2020. A New Framework for Exploring Process Controls of Flow Duration Curves. Water Resources Research, 56(1).

Ghotbi, S., Wang, D., Singh, A., Mayo, T. and Sivapalan, M., 2020. Climate and Landscape Controls of Regional Patterns of Flow Duration Curves Across the Continental United States: Statistical Approach. Water Resources Research, 56(11).

Huang, Z. et al., 2020. Differing roles of base and fast flow in ensemble seasonal streamflow forecasting: an experimental investigation. Journal of Hydrology, 0: 125272–.

Karst, N., Dralle, D. and Müller, M.F., 2019. On the Effect of Nonlinear Recessions on Low Flow Variability: Diagnostic of an Analytical Model for Annual Flow Duration Curves. Water Resources Research, 55(7): 6125-6137.

Kwakye, S.O. and Bárdossy, A., 2022. Quantification of the hydrological consequences of climate change in a typical West African catchment using flow duration curves. Journal of Water and Climate Change, 13(1): 26-42.

Leong, C. and Yokoo, Y., 2021. A step toward global-scale applicability and transferability of flow duration curve studies: A flow duration curve review (2000–2020). Journal of Hydrology, 603: 126984.

Liang, Z.M.L.G., 2019. Hydrological Analysis and CalculationHydrological Analysis and Calculation. China Water & Power Press, Beijing, China.

Luan, J., Liu, D., Lin, M. and Huang, Q., 2021. The construction of the flow duration curve and the regionalization parameters analysis in the northwest of China. Journal of Water and Climate Change, 12(6): 2639-2653.

MOHAMOUD, Y.M., 2008. Prediction of daily flow duration curves and streamflow for ungauged catchments using regional flow duration curves. Hydrological sciences journal, 53(4): 706-724.

Müller, M.F. and Thompson, S.E., 2016. Comparing statistical and process-based flow duration curve models in ungauged basins and changing rain regimes. Hydrology and Earth System Sciences, 20(2): 669-683.

Nash, J.E. and Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I — A discussion of principles. ECOLOGICAL MODELLING, 10(3): 0–290.

Nathan, R.J. and McMahon, T.A., 1990. Evaluation of automated techniques for base flow and recession analyses. Water Resources Research, 26(7): 1465 - 1473.

Over T M, F.W.H.R., 2018. Refinement of a regression-based method for prediction of flow-duration curves of daily streamflow in the conterminous United States. US Department of the Interior, US Geological Survey.

Pendergrass, A.G., 2018. What precipitation is extreme?, 360(6393): 1072–1073.

Ridolfi, E., Kumar, H. and Bárdossy, A., 2020. A methodology to estimate flow duration curves at partially ungauged basins. Hydrology and Earth System Sciences, 24(4): 2043-2060.

Saulo, H., Bourguignon, M., Zhu, X. and Balakrishnan, N., 2018. Some simple estimators for the two-parameter gamma distribution. Communications in Statistics - Simulation and Computation, 0: 1-13.

Schär, C. et al., 2016. Percentile indices for assessing changes in heavy precipitation events. Climatic Change, 137(1-2): 201-216.

SMAKHTIN, V.Y., HUGHES, D.A. and CREUSE-NAUDIN, E., 1997. Regionalization of daily flow characteristics in part of the Eastern Cape, South Africa. Hydrological sciences journal, 42(6): 919-936.

Spearman, C., 2010. The proof and measurement of association between two things. International Journal of Epidemiology, 39(5): 1137-1150.

Wang, H.J., Merz, R., Yang, S., Tarasova, L. and Basso, S., 2023. Emergence of heavy tails in streamflow distributions: the role of spatial rainfall variability. Advances in Water Resources, 171: 104359.

Wolff, W. and Duarte, S.N., 2021. Toward geostatistical unbiased predictions of flow duration curves at ungauged basins. Advances in Water Resources, 152: 103915.

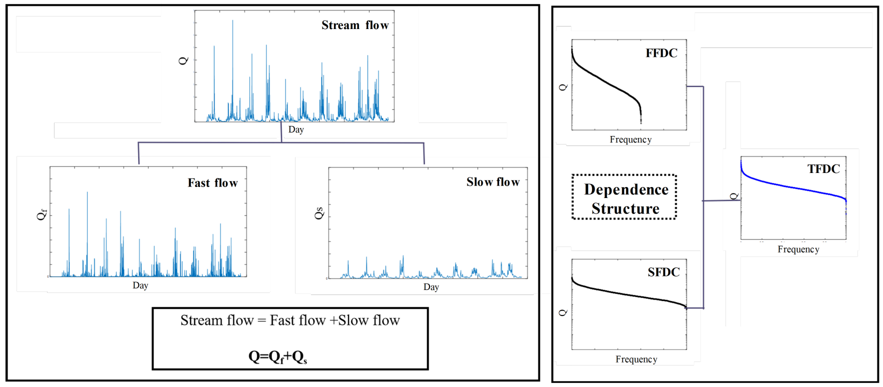
Yang, S., Gao, M., Liu, J., Wu, P. and Yang, Y., 2023. Regionalizing Streamflow Regime Function through Integrations of Geographical Controls in Mountainous Basins. Water, 15(2): 280.

Ye, S., Yaeger, M., Coopersmith, E., Cheng, L. and Sivapalan, M., 2012. Exploring the physical controls of regional patterns of flow duration curves – Part 2: Role of seasonality, the regime curve, and associated process controls. Hydrology and Earth System Sciences, 16(11): 4447-4465.

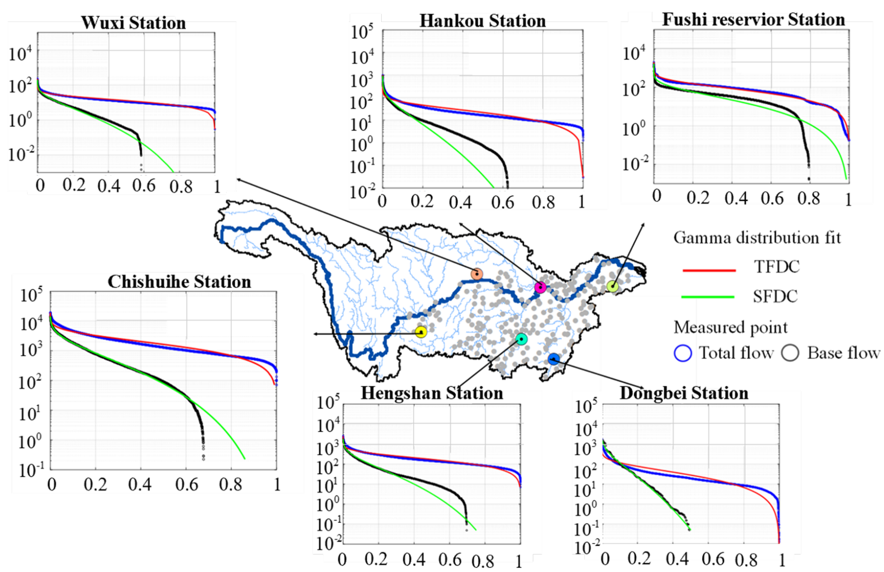
Yire Shin, J.P., 2022. Modeling climate extremes using the four-parameter kappa distribution for r-largest order statistics

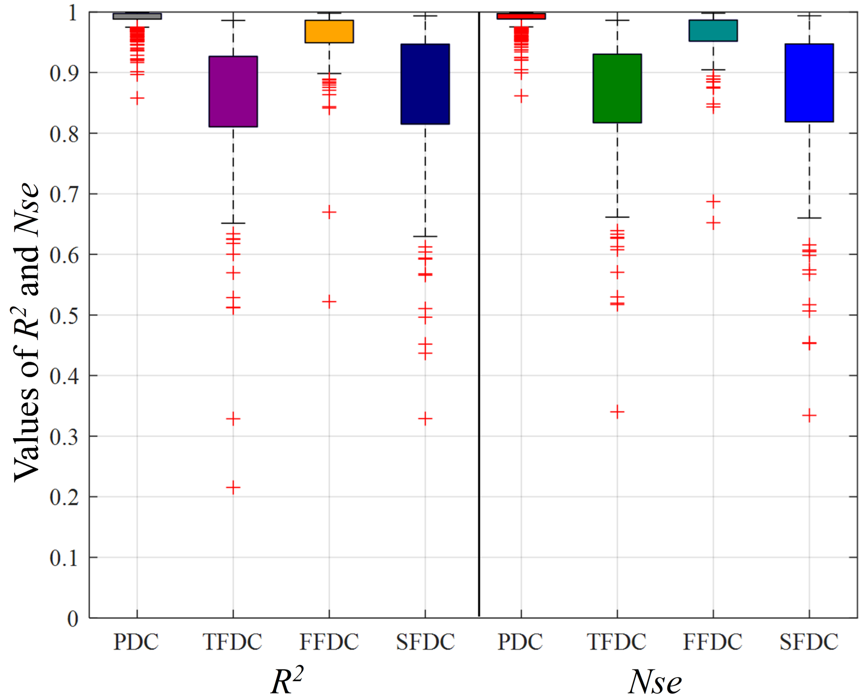
. weather climate extremes.

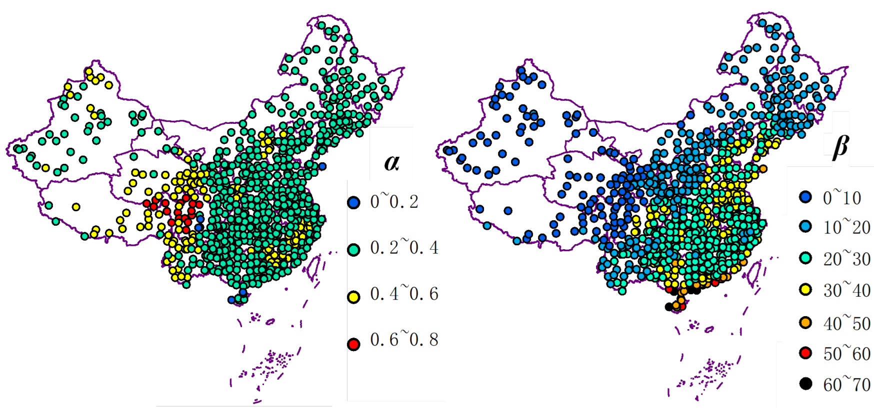
Yokoo, Y. and Sivapalan, M., 2011. Towards reconstruction of the flow duration curve: development of a conceptual framework with a physical basis. Hydrology and Earth System Sciences, 15(9): 2805-2819.



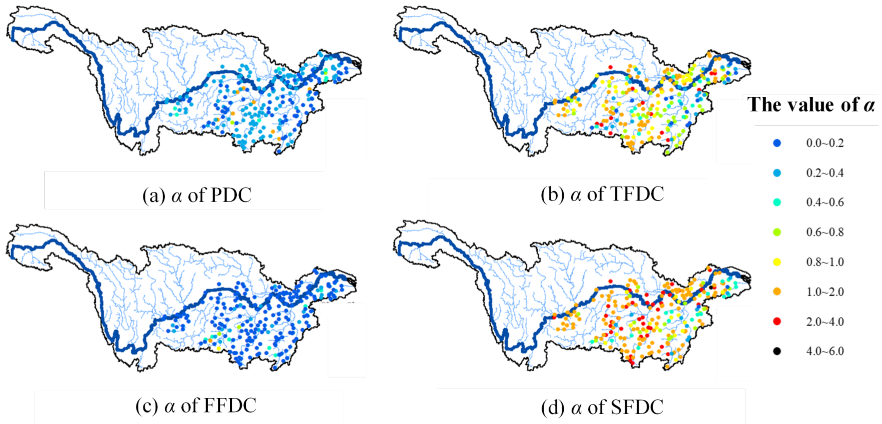
(a) Time series (b) Probability distribution



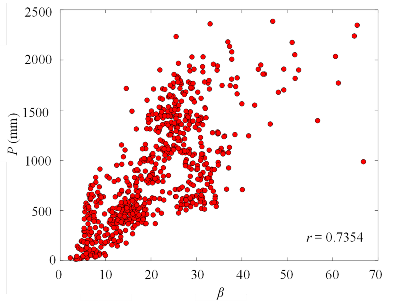




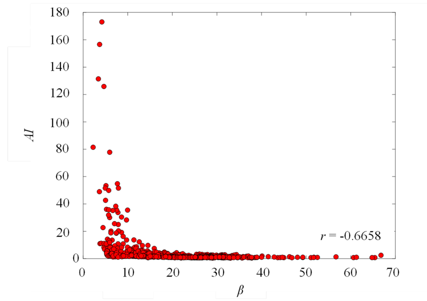
(a) Spatial distribution of *α* **(b) Spatial distribution of *β***



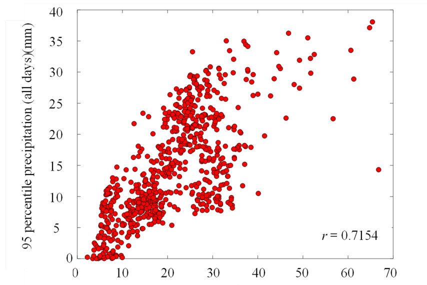
(a) α of PDC
(b) α of TFDC
(c) α of FFDC
(d) α of SFDC



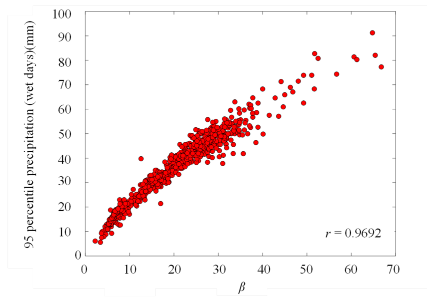
(a) β and precipitation



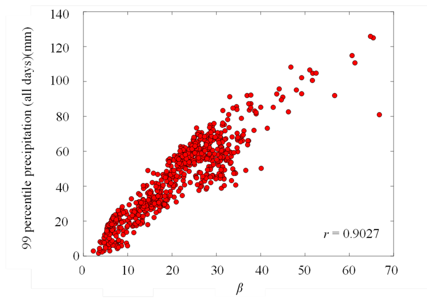
(b) β and aridity index



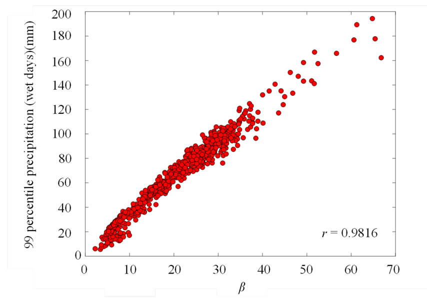
(c) β and 95 percentile precipitation (all days)



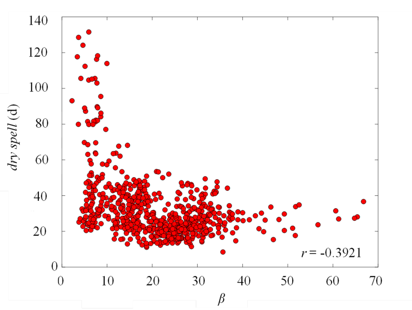
(d) β and 95 percentile precipitation (wet days)



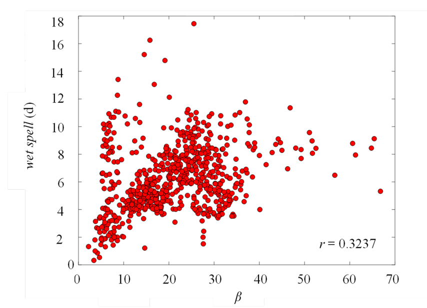
(e) β and 99 percentile precipitation (all days)



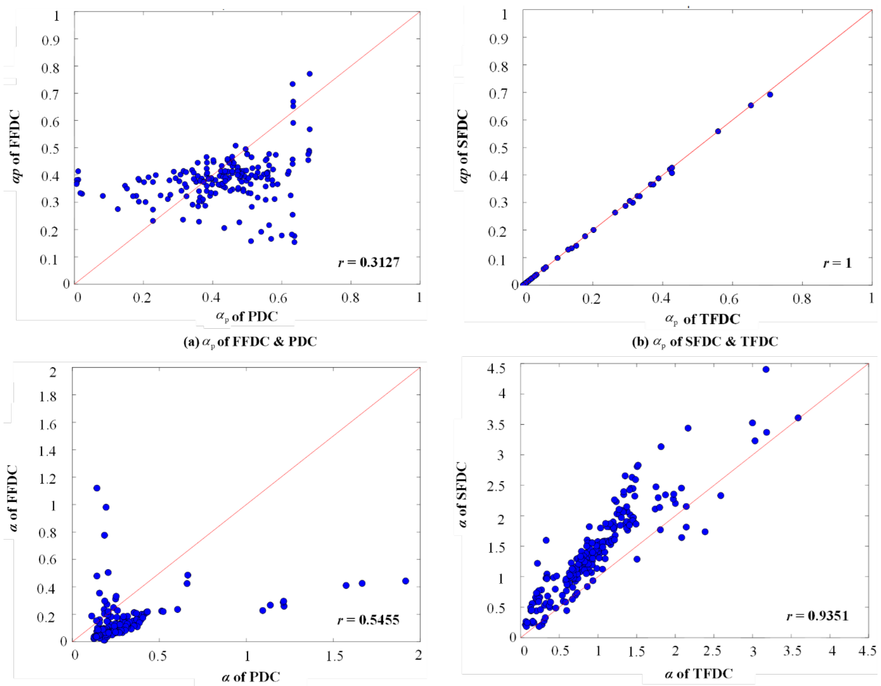
(f) β and 99 percentile precipitation (wet days)



(g) β and *dry spell*



(h) β and *wet spell*

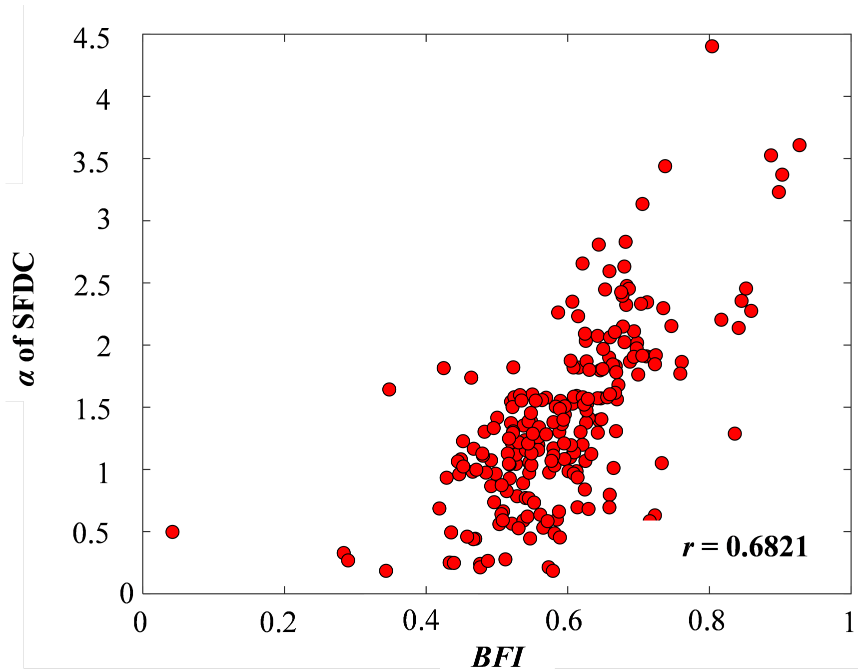


(a) α_p of FFDC & PDC

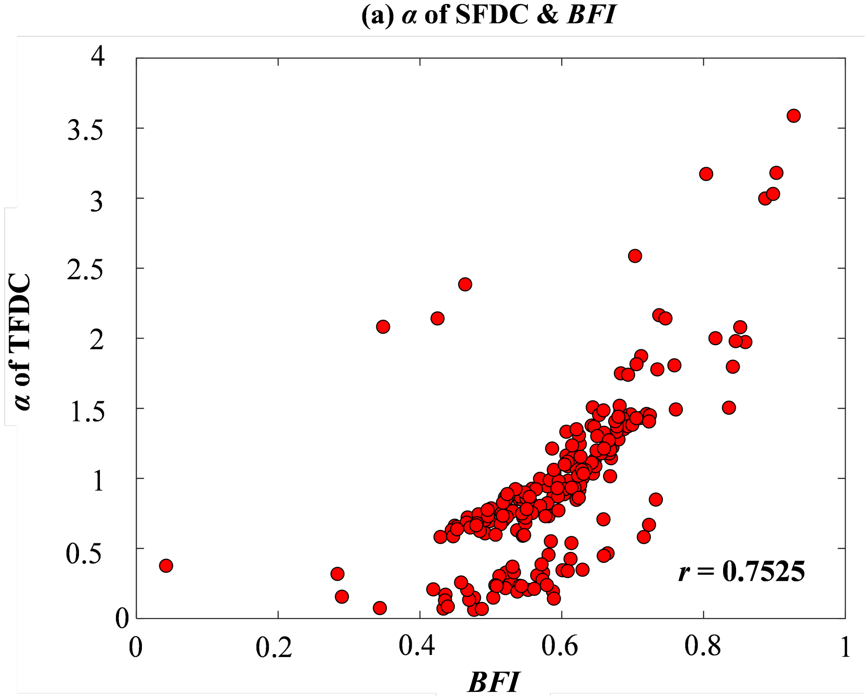
(b) α_p of SFDC & TFDC

(c) α of FFDC & PDC

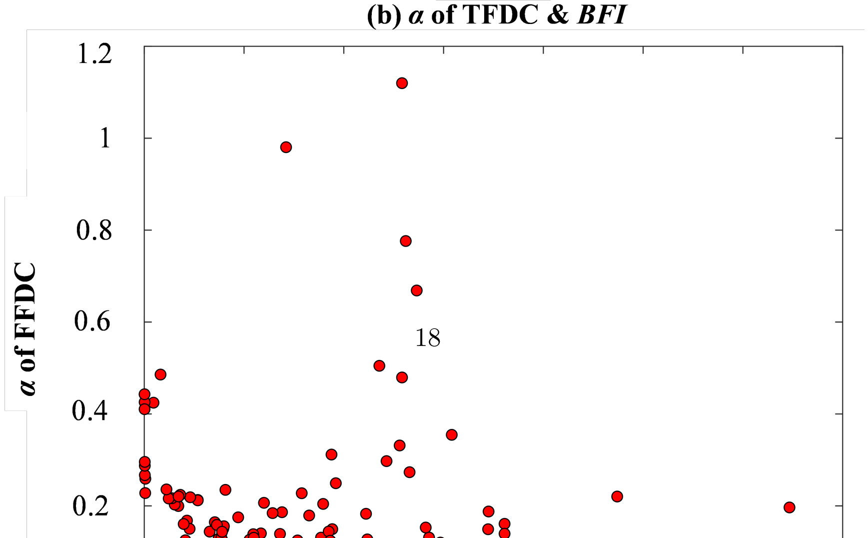
(d) α of SFDC & TFDC

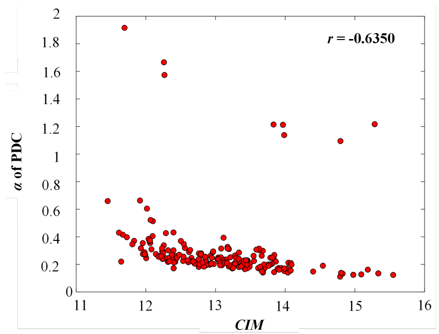


(a) *α* of SFDC & *BFI*

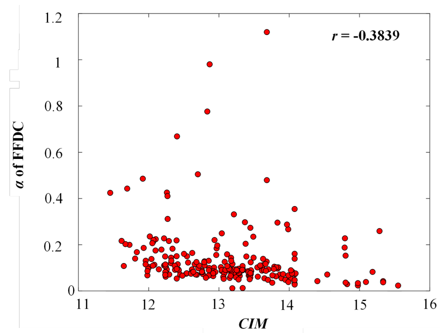


(b) *α* of TFDC & *BFI*





(a) α of PDC & CIM



(b) α of FFDC & CIM