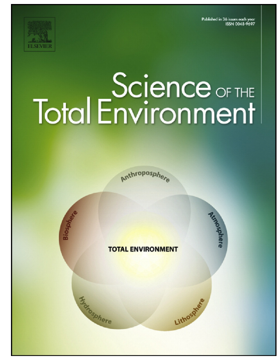
Journal Pre-proof



Control of climate and physiography on runoff response behavior through use of catchment classification and machine learning

Shuping Du, Shanhu Jiang, Liliang Ren, Shanshui Yuan, Xiaoli Yang, Yi Liu, Xinglong Gong, Chong-Yu Xu

PII: S0048-9697(23)05047-7

DOI: https://doi.org/10.1016/j.scitotenv.2023.166422

Reference: STOTEN 166422

To appear in: *Science of the Total Environment*

Received date: 16 June 2023

Revised date: 17 August 2023

Accepted date: 17 August 2023

Please cite this article as: S. Du, S. Jiang, L. Ren, et al., Control of climate and physiography on runoff response behavior through use of catchment classification and machine learning, *Science of the Total Environment* (2023), https://doi.org/10.1016/j.scitotenv.2023.166422

Control of climate and physiography on runoff response behavior through use of catchment classification and machine learning

Shuping Du[a], Shanhu Jiang[a, b*], Liliang Ren[a,b], Shanshui Yuan[a] , Xiaoli Yang [a,], Yi Liu [a,], Xinglong Gong[c], Chong-Yu Xu[d]

[a]*College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China*

[b]*The National Key Laboratory of Water Disaster Prevention, Hohai University, Nanjing 210098, China*

[c]*School of Water Conservancy & Civil Engineering, Northeast Agricultural University, 150030, Harbin*

[d]*Department of Geosciences, University of Oslo, Oslo, Norway*

*Corresponding author:

Prof. Shanhu Jiang

State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210098, China for Prof.

Prof. Shanhu Jiang *Email: hik0216l@hhu.edu.cn*

Abstract

Understanding of runoff response changes (RRC) is essential for water resource management decisions. However, there is a limited understanding of the effects of climate and landscape properties on RRC behavior. This study explored RRC behavior across controls and predictability in 1003 catchments in the contiguous United States (CONUS) using catchment classification and machine learning. Over 1000+ catchments are grouped into ten classes with similar hydrological behavior across

CONUS. Indices quantifying RRC were constructed and then predicted within each class of the 10 classes and over the entire1000+ catchments using two machine learning models (random forest and CUBIST) based on 56 indicators of catchment attributes (CA) and 16 flow signatures (FS). This enabled the ranking of the important influential factors on RRC. We found that (i) CA/FS-based clusters followed the ecoregions over CONUS, and the impact of climate on RRC seemed to overlap with physiographic attributes; (ii) CUBIST outperforms the random forest model both within the cluster and over the whole domain, with a mean improvement of 39% (depending on clusters) within clusters. Runoff sensitivity was better predicted than runoff changes; (iii) FS related to runoff ratio, average, and high flow are the most important for RRC, whereas climate (evaporation and aridity) is a secondary factor; and (iv) RRC patterns are substantial in the dominant factor space. High total changes and catchment characteristic-induced changes occurred mainly at 100°west longitude. The elasticity of climate and catchment characteristics was found to be high in spaces with high evaporation and low runoff ratios and low in spaces with low evaporation and high runoff ratios. Uncertainties existed in the number of catchments between clusters which was verified using a fuzzy clustering algorithm. We recommend that future research that clarifies the impact of uncertainty on hydrological or catchment behavior should be conducted.

Keywords: hydrological behavior; climate change; dominant controls; random forests; CUBIST; uncertainty analysis

1. Introduction

Water resources are facing significant changes due to climate alteration and anthropogenic activities (Shao et al. 2020, Maftouh et al. 2023). Estimates of this change are essential for supporting policies and planning decisions. Understanding the hydrological response to climate change/human modifications is one of the ways to accomplish this aim. A lot of recent studies have focused on the spatial patterns of hydrological responses, particularly runoff responses (Wu et al. 2017, Yi et al. 2017, Wang et al. 2018, Forbes et al. 2019). However, there is little understanding of the physical significance behind the spatial patterns of hydrological responses, and a clear vision of this would be helpful for developing hydrological models and make decisions in water resource management (Andreassian et al. 2016; Wang &Ni, 2023; Wang et al.,2023).

Runoff is a sensitive component of hydrological recycling (Balha et al., 2023). Runoff sensitivity and its changes due to climate change and/or human activities are referred to as runoff response changes (RRC). This is usually used to capture past changes in runoff behavior and for evaluation of estimating future water resources (Zhang et al. 2008, Wang and Hejazi 2011; Abra et al.,2023;Wang et al.,2023). Numerous studies have focused on the spatial variability of the RRC across region, basin, and globally (Shahid et al.,2021; Bajirao et al.,2022; Bharat & Mishra,2021). For example, Wang and Hejazi (2011) used a decomposition method based on the Budyko framework (Budyko,1974) to quantify the impact of precipitation, potential evaporation, and direct human activity on mean annual streamflow over 413 watersheds in the contiguous

United States (CONUS). Li et al. (2022) employed a statistical regression method to investigate the variations in streamflow in the Jialing River Basin, China. At the beginning of this century, Sankarasubramanian et al. (2001) explored the pattern of precipitation elasticity across the CONUS, revealing a contour map of precipitation elasticity that tended to be high for basins with water and radiation inputs seasonally out of phase with one another. Recently, Gong et al. (2022) built a relationship between the RRC and hydrometeorological factors (e.g., annual rainfall and aridity index) over the CONUS. However, there is still a limited understanding of why RRC presents a unique pattern and the physical reasons behind it. Understanding RRC behavior related to dominant factors and predictability is important for exploring the physical mechanisms behind RRC spatial patterns (Yin et al.,2023; Slater et al.,2023; Fang & Pomeroy,2023). Capturing future changes in streamflow in the context of climate change, in addition to correcting the bias of the runoff-rainfall model, also requires an understanding of RRC behavior (Andréassian et al. 2016, Di et al. 2017;Gao et al.,2019; Ghamariadyan & Imteaz 2021).

Each catchment is a synthetic system driven by a combination of climate and landscape, which makes it difficult to identify and capture the hydrological behavior within a catchment (Addor et al. 2017; Blöschl et al., 2019; Beevers et al.,2021). Catchment classification is a useful tool for exploring the laws and patterns of this behavior within a catchment (Sivapalan, 2003; Cinkus et al.,2023). Generally, catchments are grouped based on measures of similar seasonality, such as physiographic and meteorological characteristics; however, these measures tend to be weak for large-scale catchments

(Burn, 1997; Wagener et al., 2007). Recently, datasets comprising a large number of catchments and hundreds of attributes have become available worldwide, enabling the capture of new insights into hydrological behavior (Addor et al. 2017, Alvarez-Garreton et al. 2018; Newman et al., 2014; O'Sullivan et al.,2023). Jehn et al. (2020) used eight hydrological signatures to classify more than 600 catchments from the CAMELS dataset over the CONUS. To explore the physical controls of the spatial patterns of flow signatures (FSs). Kuentz et al. (2017) used sixteen FSs and thirty-five catchment attributes (CAs) to classify over 30000 catchments in Europe. To the best of our knowledge, no large sample study has used CAs and FSs with good predictability to explore the spatial patterns of the RRC based on method of catchment classification. In addition, the difference in the number of catchments between classes cannot be reduced using traditional classification methods (Kuentz et al. 2017), and there is little knowledge about how this uncertainty in the number of catchments impacts hydrological behavior.

Machine learning (ML) method is robust and performs well in prediction, data mining, and model parameterization in hydrological field (Shen, 2018; Zhao et al., 2019; Ma et al., 2021; Tsai et al., 2021). ML has multiple advantages in data analysis: (i) it allows multiple predictors and non-linear relationships; (ii) there is no need to know the underlying mechanism between the input and output; and (iii) there is no problem of data overfitting. Addor et al. (2018) employed random forest to explore the predictability of signatures in more than 600 catchments over the CONUS. Cheng et al. (2022) used a boosted regression tree to model the partitioning of precipitation into runoff and evaporation based on the Budyko framework. As most of MLs are "black box model", they are weak in straightforwardly quantifying the underlying relationships between inputs and outputs. Fortunately, interpretable artificial intelligence can overcome this problem (Basagaoglu et al., 2022). The accumulated local effect (ALE), a robust and advanced global diagnostic algorithm, can rank features based on dependent variables (Gao et al.,2023; Apley et al.,2020). This can thus let modeler identify dominant factors of dependent variables and influence the process between predictors and dependent variable, thus visualizing "black model" clearly. Most previous studies focusing on the identification of the dominant factors of hydrological variables in hydrology modelling tend to use regression trees that consider linear controls (Padrón et al., 2017; Zhta et al.,2023; Basri et al.,2023), and only a few studies have interpreted the interior relationship between inputs and outputs while considering non-linear relationships (Bai et al. 2019; Chang et al.,2023; Haddad& Rahman,2023). In addition to the dominant factors, it is necessary to investigate the predictability of RRC using the ML method as the need for water resource management in ungauged basins, especially to understand whether or how catchment classification with similar catchment behavior improves the predictive skill of the ML model.

In this study, catchment classification and two MLs (random forest and CUBIST) were used to explore the physical mechanisms of the RRC behavior. Firstly, 1003 catchments across the CONUS were grouped into 10 classes based on 56 CAs in terms of climate, topography, geology, hydrology, and soil properties, and 16 FSs with good predictability in space. Thereafter, two MLs were used to investigate the predictability of the RRC, and subsequently, ALE was used to identify the dominant factors considering the non-linear relationship over the entire domain and within each cluster. The primary goals of present study were to: (i) classify catchments of GAGE-II dataset based on 56 CAs and 16 FSs, respectively, and investigate catchment behavior within each cluster; (ii) to predict RRC within each cluster and over whole domain, and understand how this catchment classification determine predictability of RRC compared to modelling over whole domain where all catchments as a input; (iii) identify which of 56 CAs and 16 FSs were the key factors controlling RRC; and (iv) explore what patterns of RRC is in the space of dominant factors. Compared with previous studies related to RRC, our novelty is that the combination of catchment classification and ML was used to explore RRC behavior (dominant controls and predictability) with consideration of non-linear relationships over large scale domains.

2. Methodology and Data

2.1. Budyko Framework

The Budyko framework is a simple computational formulation that relates the ratio between actual evaporation and precipitation [*E/P*] to the ratio between potential evaporation and precipitation [*PE/P*], and the shape parameter representing catchment properties (Budyko, 1958). The equation of Yang et al. (2008), a commonly used Budyko-type form, uses parameter n to represent the shape parameter. This form is expressed in Eq. (1):

$$E = \frac{PE \times P}{(PE^n + P^n)^{1/n}} \quad (1)$$

where E is the long-term actual evapotranspiration (mm), P is the long-term average annual precipitation (mm), *PE* is the long-term average potential evaporation (mm), and *n* is the catchment characteristic that represents the interaction between climate, soil, and plants.

If Budyko's equation is defined as *f=(PE, P, n)*, its differential form is expressed as follows:

$$dE = \frac{\partial f}{\partial P}dP + \frac{\partial f}{\partial PE}dPE + \frac{\partial f}{\partial n}dn \tag{2}$$

In the long term, the change in the average soil water storage was approximately zero for catchment areas >1000 km^2 and long-time series (>11 years). Therefore, Eq. (2) can be rewritten as follows:

$$dQ = \left(1 - \frac{\partial f}{\partial P}dP\right)dP - \frac{\partial f}{\partial PE}dPE - \frac{\partial f}{\partial n}dn \tag{3}$$

where *Q* is the average runoff depth (mm), which is the long-term average value from 1941-2015 in the present study.

By dividing the left and right sides of Eq. (3) using Q, the relative runoff change (%) can be obtained as follows:

$$\frac{dQ}{Q} = \varepsilon_p\frac{dP}{P} + \varepsilon_{PE}\frac{dPE}{PE} + \varepsilon_n\frac{dn}{n} \tag{4}$$

Where ε_p, ε_{PE}, ε_n is the precipitation, potential evaporation, and catchment characteristic elasticity respectively, that is, runoff sensitiveness in this study.

A non-parametric method was used to estimate the temperature elasticity, considering the large scale of the study area and the lack of meteorological series in terms of relative humidity, wind speed, and radiation. This method has been proven to have low bias and is non-parametric (Yang et al., 2008). Its form is expressed by Eq. (5):

$$\varepsilon = \frac{\bar{X}}{\bar{Q}} \frac{\sum(X_i-\bar{X})(Q_i-\bar{Q})}{\sum(X_i-\bar{X})^2} \tag{5}$$

where X_i, Q_i is climatic variable values and runoff depth in the i-th year, respectively, and ε is climate elasticity.

The potential evaporation was estimated using the Penman-Monteith temperature (PMT) method (Moratiel et al., 2020). Although there is a deviation from the results calculated using the Penman-Monteith FAO-56 equation, some scholars have shown that the deviation is negligible, particularly for potential evaporation on an annual scale (Moratiel et al., 2020; Senatore et al., 2020).

2.2 Catchments classification

We used a hierarchical minimum-variance clustering method to group catchments, as done by Kuentz et al. (2017) and Li et al. (2018), which combined the k-means algorithm (Hartigan and Wong, 1979) and Ward's minimum-variance method (Ward, 1963). Kuentz et al. (2017) suggested that 10 or 11 is a reasonable number of clusters using the elbow and two other methods who grouped more than 35000 catchments over Europe based on catchment descriptors. Furthermore, Jehn et al. (2020) recommended that 10 clusters are more appropriate than the CONUS. Therefore, we used ten clusters for our dataset without validation. Before the classification, we used principal component analysis to reduce the problem of information correlation for the flow signatures. This enabled the resulting FSs to account for at least 60% of the total variance in the original flow signatures. As for the 56 CAs, there is no significant correlation between them (Figure S8 in Supporting Information S1), and CAs cover varied types of catchment properties; principal components cannot comprehensively represent them; thus, we use all of them to classify our catchments.

2.3 Machine Learning

2.3.1 Random Forest

A random forest, which comprises an ensemble of regression trees, can predict dependent variables as regression tools and classify discrete variables as classification machines (Pal, 2005). This method is commonly used in various fields of geoscience and engineering (Booker & Woods, 2014; Chaney et al., 2016). Breiman (2001) provides detailed description for the method. We used this machine learning model for data mining of a large-scale set for five reasons. These included, (i) allowing multiple predictors and non-linear relationships. Hydrological behavior is usually determined by multiple catchment attributes, and a single relationship between the RRC and catchment attributes tends to be weak (Beck et al.,2015). In addition, the hydrological response to catchment attributes tended to be non-linear. Random forests are more useful for the above task because they allow for multiple predictors and non-linear relationships because they consist of a series of thresholds compared to classical regression models. (ii) There is no need to know the underlying mechanism between the input and output. Random forests can reveal the relationship without the assumption of the physical principles of hydrological processes, and the output can be explained well by the interpretable machine learning method. (iii) Reduced risk of data overfitting. The ensemble of regression trees of random forests enables predictions to be absent from the influence of specific predictors; and (iv) transparency and interpretability in the importance of predictors. Random forests can rank the importance of predictors on response variables, and lastly (v) perform well in the prediction and provide reliable uncertainty estimates.

We used 500 trees to predict each RRC based on 56 CAs and 16 FSs and allowed each tree to grow freely, making random forests very robust. Generally, 500 trees would produce a good prediction (Addor et al., 2018). The predictions were evaluated through ten-cross-validation, which also determined the number of variables at each split (mtry). We used 70% of the catchments as the training set and the remainder as the testing set. This procedure was repeated nine times. The accuracy of the model was evaluated by determining the coefficient of determination (R^2). All procedures were completed using the “randomForest” package in R (James et al., 2013).

2.3.2 CUBIST

Traditional classification and regression trees (CART) use the average of all training values at the final leaf node for sample prediction, which leads to a significant bias for the new sample set and makes it difficult to achieve ideal prediction results (Breiman,1984). CUBIST is a segmented linear tree model developed from the M5 model tree (Quinlan, 1992; Viscarra et al., 2019), that can efficiently overcome the disadvantages of CART. The advantage of CUBIST is that the leaf node of the tree is a linear regression model, and a series of segmented linear models are combined to form a CUBIST regression tree that can effectively solve problems related to non-linearity. In addition, the training rules of CUBIST are simple, effective, and fast, and the segmentation of the input space is automatically performed by an algorithm that can handle high-dimensional problems (Time et al.,2016).

CUBIST divides the input sample space into rectangular areas with parallel edges. In each layer of the model tree, the attribute with the strongest recognition power acts as the root node of the subtree, and the samples are divided into several subsets according to the root node (J.F.A.,1914). To prevent excessive growth, multiple constraints are set for node growth: the ratio or difference in the standard deviation of the target attribute between the node sample and population sample is less than a certain threshold, or the number of samples of the node is less than a certain threshold (Fernández-Delgado, et al.,2018). After establishing the preliminary model tree, it is necessary to prune the tree, that is, merge some miscellaneous subtrees and replace them with leaf nodes to improve the efficiency and conciseness of CUBIST (He et al.,2021). Finally, smoothing methods must be used to compensate for the discontinuous rows of leaf nodes after pruning.

CUBIST optimizes the modelling by setting the number of rules (0-100) and instances (0-9). The number of rules refer to the number of models that must be used in the tree model, and the instances refer to the number of samples that must be referenced to correct the results during forecasting. A higher number of rules does not imply a higher accuracy of the model (Khaledian & Miller,2020). To avoid overfitting and improve the stability and accuracy of the model, parameter optimization is necessary before training the model. The parameters in CUBIST were obtained through ten-fold cross-validation by minimizing the Mean Square Error (MSE) (Viscarra et al., 2019). In the present study, the maximum number of rules was set to 5, the extrapolation factor was 10, and the committee models (1,2, 3), and the number of neighbouring observations (1, 2) were different for each RRC. In CUBIST, feature importance can be evaluated by the percentage of models in which each variable is used, that is, the usage of variables [%].

2.4 Data

The CA and daily discharge data in this study were obtained from the GAGES-II dataset proposed by the United States Geological Survey (USGS; http://esapubs.org/Archive/ecol/E091/045/default.htm)(Falcone et al., 2011). Geospatial catchment data encompass climate, hydrology, landform, land use/cover, geology, and human activities (Falcone et al., 2011). In our study, six catchment attribute types were considered, including topography, geology, climate, hydrology, land use/cover, and soil characteristics (the list of CA, details on methods, and data sources used to estimate them are provided in Table S1 in Supporting Information S1). Topographic features such as mean elevation and slope were extracted from the USGS database; geological features such as the proportion of geological types and catchment area were obtained from the USGS; climatic characteristics such as mean precipitation (rainfall and snow) and aridity were computed using hydrometeorological daily time series from Parameter-Elevation Regressions on Independent Slopes Model (PRISM); hydrological characteristics (baseflow ratio, topographic wetness index, etc.) were obtained from the corresponding database; land use/land cover such as fraction of forest and developed area were retrieved from U.S. Department of Agriculture (USDA); soil characteristics such as average permeability and average value of bulk density were extracted from USGS. These catchment attributes were selected to quantitatively describe the impact of landscape characteristics on hydrological processes as much as possible (Addor et al., 2017a, 2017b). Daily discharge at the GAGES-II hydrological sites was obtained from the USGS Water Information System (https://waterdata.usgs.gov/nwis/). Completeness and length of discharge of at least 20 years, and 80% completeness were considered for daily discharge (Falcone et al., 2011). We considered the period from 1941 to 2015 as the study period, and sites with a runoff ratio less than one, and8 finally selected 1003 sites for our study. FS were selected to represent the different characteristics of hydrographs across different time scales and have been widely used to quantify the sensitivity to hydrological processes occurring across scales. FSs were estimated by daily discharge data, and details on the methods and descriptions have been provided in Table S1 in Supporting Information S1. The criteria for selecting the FSs followed those of Westerberg and McMillan (2015). Generally, an FS should satisfy the requirements of determination, robustness, consistency, representativeness, and discrimination. In this study, 16 FSs were selected on the basis of these criteria. Maps of watershed boundaries were obtained from the U.S. Department of Agriculture-Natural Resources Conservation Service, USDA-NRCS) (https://www.nrcs.usda.gov/wps/portal/nrcs/main/national/water/watersheds/). Meteorological time series including daily precipitation as well as maximum and minimum daily temperatures were obtained from the PRISM dataset (PRISM Climate Group, 2016). The original grid resolution of the PRISM dataset was 800 m, but it was filtered to a 4 km resolution for ease of download and manipulation. The framework of this study is illustrated in Figure1.

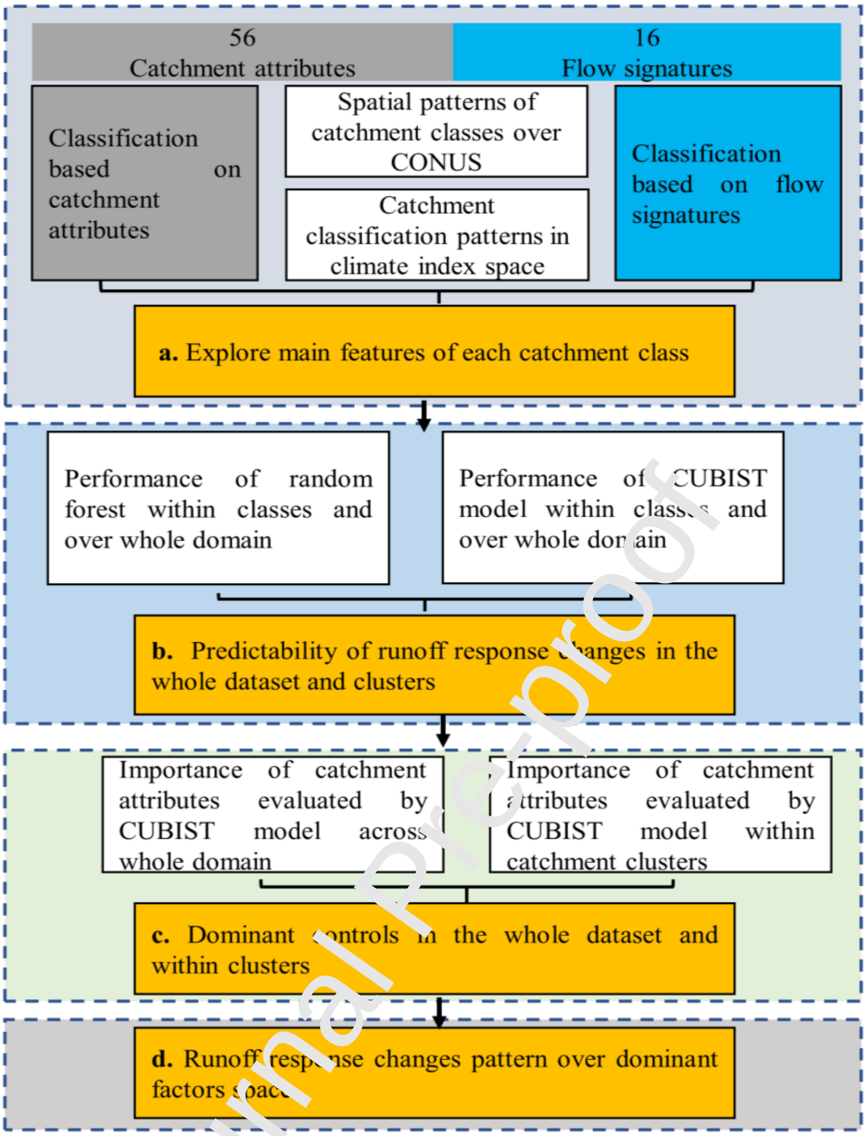


Figure 1. Framework of this study.

3. Results

3.1. Main Features of Catchment Classification

Catchments located within one cluster could be very far apart in space (Fig.2); for example, the distance between catchments in cluster 5 CA-based and FS-based catchments is up to 4000 km, which is almost the north-south distance of CONUS (Fig.2c). This implies that varied climatic conditions may produce similar catchment behavior represented by CA and FS (Fig.2d), also demonstrates the importance of regional studies in hydrological field (McCabe and Wolock 2014). Furthermore, the influence of human disturbance can also result in substantial alterations in catchment behavior (Alvarez-Garreton et al. 2018, Zhang et al. 2021). Hydrological modelling and data driven conceptual frameworks have been commonly used to separate climate change from anthropogenic influences that affect the water balance (Yang et al.,2008; Gong et al.,2022; Zhang et al.,2023; Nury et al., 2023). Some catchments in our study have suffered disturbances from human activities, such as dam construction (Falcone et al., 2011). Interestingly, the spatial pattern of clusters across ecoregions over the CONUS can enable the separation of the effect of climate change on catchment behavior from human activities as unique catchment behavior within each class, and relatively little human disturbance over ecoregions (Dennison et al. 2014, Auch et al. 2016, Jehn et al. 2020). Hydrological behavior within clusters over ecoregions could be considered to be impacted solely by catchment behavior without human influence.

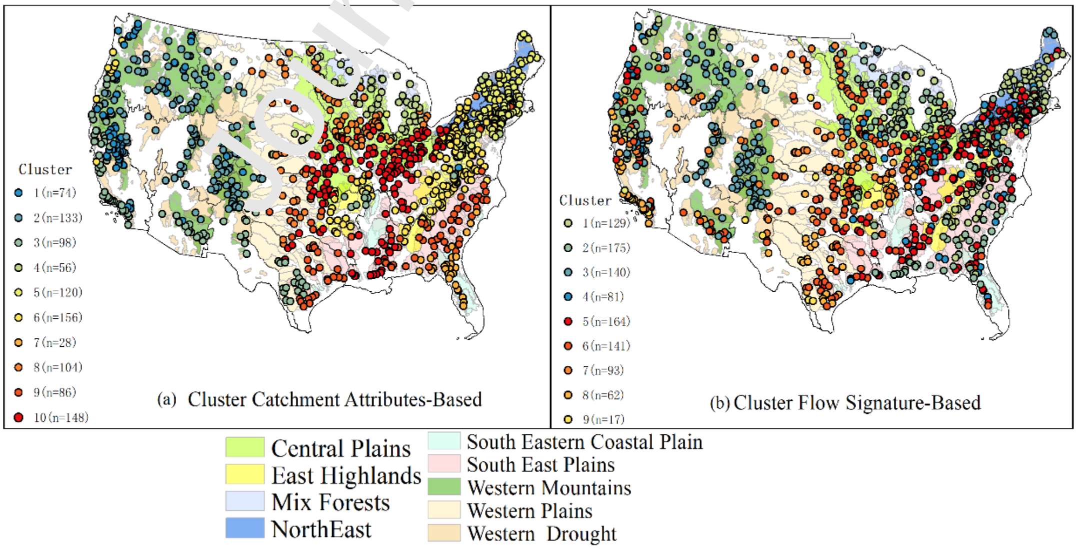


Figure 2. Spatial patterns of catchment clusters over Continental U.S. based on (a) catchment attribute (CA); (b) flow signature (FS).

We summarized the spatial patterns of each class and typical features within each class (Table 1), box plots of the features in each CA/FS-based cluster are shown in Figures S1 and S2 in Supporting Information S1. We can infer that catchment classification based on CA and/or FS can improve the understanding of catchment and hydrological behaviors captured by CA and FS. Catchment behavior is mainly controlled by climate. For instance, cluster 1 CA-based (7%) is characterized by a low topographic wetness index, which is mainly in the western mountains where heavy rainfall has occurred in winter which falls as ice and snow; at the same time, the region has a higher summer runoff and runoff ratio as large amounts of snow melts to generate runoff in late spring and early summer (Rice et al. 2016). Cluster 4 CA-based (6%), however, is characterized by high topographic wetness, which is mainly in southeast plains where there is high annual rainfall (> 1000 mm per year on average), uniform distribution, and high potential evaporation. Meanwhile, climate also control the flow signature behavior, the FS-based cluster 1 (13%) has the typical feature of an earlier mean half-flow date, which is distributed in the northeast and southeast plains. Cluster 3 FS-based is, which is represented by a later mean half-flow date, in the western mountains. Furthermore, landscape surface processes such as vegetation and land use/land cover also dominate the catchment/hydrological behavior (Table 1); CA-based clusters 5-6 in the northeast have characteristics of low bulk density and high soil material compared to CA-based cluster 8 in western regions of CONUS; FS-based clusters 1-2 in the southeast are represented by earlier mean half-flow-date and later half-flow-date; however, compared to cluster 3 in western regions.

Table 1. Typical features of each cluster. Catchments are classified based on catchments attributes (CA) and flow signature (FS) respectively. Typical features in each cluster are defined by the lowest coefficient of variation among all features within the cluster.

Cluster	Number for CA classification	Number for FS classification	Main region for CA classification	Main region for FS classification	Typical feature for CA classification	Typical feature for FS classification
1	74	129	Western mountains	Northeast and south east plains	Low topograph-ic wetness index	Earlier Mean half flow-date
2	133	175	Western mountains	Central plains and south east plains	High depth to seasonally high water table	Earlier Mean half flow-date
3	98	14[illegible]	Western drought	Western mountains	High depth to seasonally high water table	Later Mean half flow-date
4	56	81	South east plains	Mixed regions	High topograph-ic wetness index	High runoff ratio

5	12	164	Northeast	Northeast and southeast plains	Low bulk density	Earlier Mean half flow-date
6	156	141	Northeast and east highland	Central plains and western drought	Low bulk density	Later Mean half flow-date
7	28	93	South east coastal plain	Central plains and western plains	High total soil thickness	Later Mean half flow-date
8	104	62	Western plains	Western plains	High soil material with < 3 inches	Earlier Mean half flow-date
9	86	17	South east plains	Western plains	Low topograph-ic wetness index	Low flow
10	148		Central plain and south east plains	-	High bulk density	-

A clear split of the clusters along the center of the CONUS is shown in Figure2. Generally, clusters 1-4 CA-based were in the west, and the others are in the east. FS-based clusters 5-10, were mainly in the east, whereas the remaining clusters were in the west.

We further explored the cluster pattern in climate index space, as Knoben et al. (2018), using three dimensionless indices of climate in terms of aridity, seasonality, and the fraction of precipitation falling as snow to explore climate controls. As shown in Figure 3, the resulting clusters, both CS-based and FS-based, did not appear to be a split of classes in space, as observed by Knoben et al. (2018). Jehn et al. (2020) related this to the classification approach and further inferred that either the climate is overlapped by other catchment attributes, or that dynamic changes in climate cannot be adequately captured in a single catchment. In fact, neither catchment behavior nor hydrological behavior is simply determined by a single factor, they seem to be climatic originally, but are shaped through landscapes (Carmona et al. 2014, Vano et al. 2015, Stein et al. 2021). For example, FS-based cluster 9 is characterized by low flow, which is usually affected by a combination of climate, topography, soil, and geology (Kam and Sheffield 2015, Floriancic et al. 2018, Li et al. 2018). Although the classification methodology can improve the understanding of catchment/flow behavior through similarity theory, this does not mean that there is only one catchment attribute/flow signature within a cluster, as shown in Figure7 and Figure S7 in Supporting Information S1, in fact catchments within a single cluster were controlled by many aspects of influence, and we only considered the dominant aspect within the cluster.

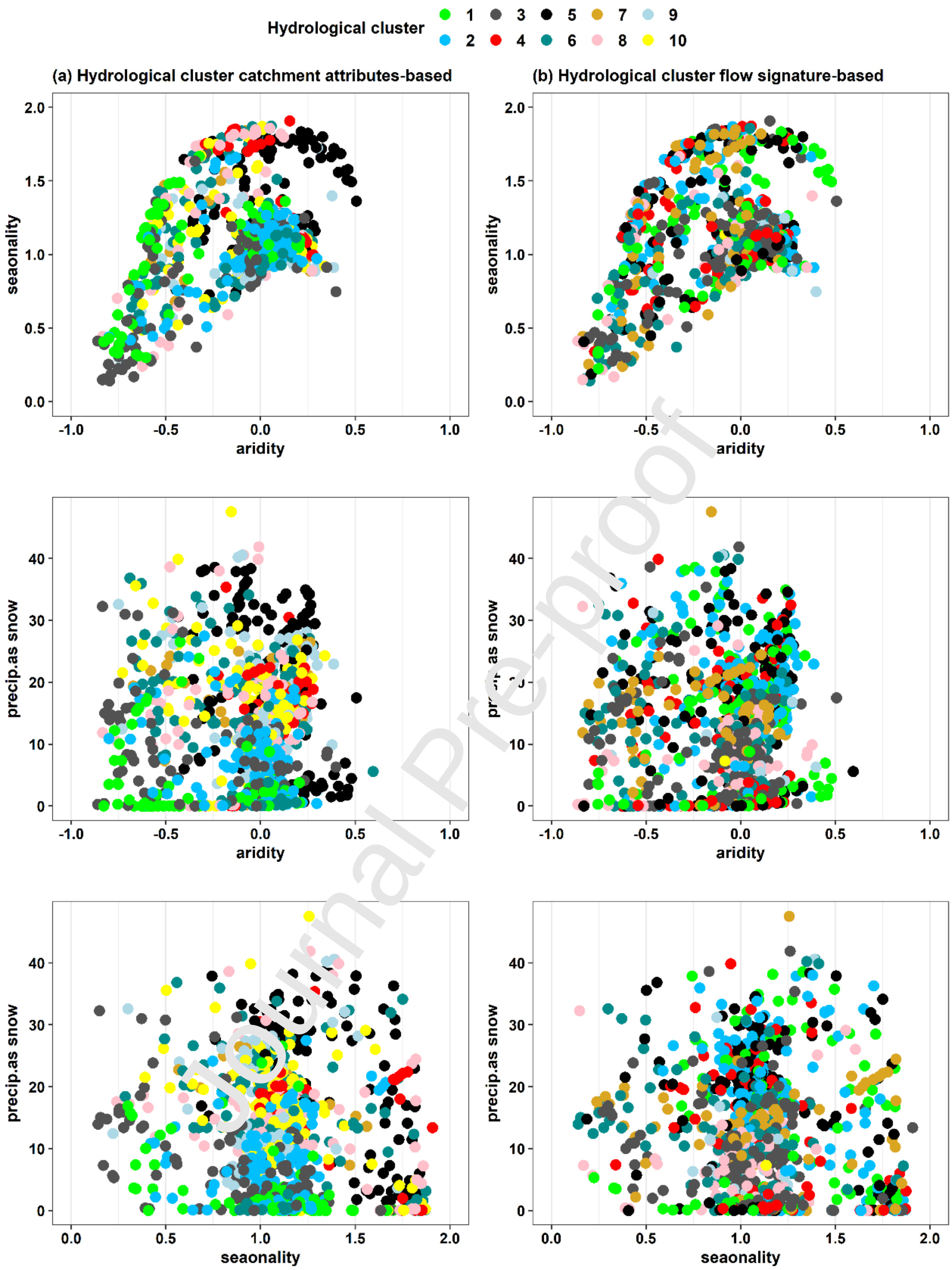


Figure 3. Clusters in climate index space based on (a) CA, (b) FS. Single dots represent catchments and are colored by hydrological cluster. There is no linkage between the numbering of the catchment classes for two classifications.

3.2. Predictability of runoff response changes in the whole dataset and within clusters

The ability to accurately predict hydrological behavior is the key for future risk evaluation and policymaking, which is a severe test (Crawford and Linsley,1966). Most studies referring to the ability to predict hydrological behavior have shown that ML has a better skill than conceptual methods to do so (Addor et al. 2018). In this study, we used two machine learning models (random forest and CUBIST) to predict the RRC in all catchments over whole domain and within each cluster, and the performance of the models was evaluated by determining the coefficient R^2. The CUBIST model performed better in predicting most RRCs across the entire dataset than random forests (apart from precipitation and potential evaporation induced changes), especially for total changes (0.36 of random forest vs 0.66 of CUBIST), catchment characteristics-induced changes (0.38 vs 0.65), and catchment characteristic elasticity (0.53 vs 0.94). Runoff sensitivity can be better predicted than runoff changes, which is expected because it is relatively independent of the influence of forcing in a single basin (Sankarasubramanian et al. 2001). We found that RRC affected by a combination of factors tended to be better predicted. For example, the performance of CUBIST for total changes and catchment characteristic elasticity had the highest R^2 with 0.66 and 0.94 respectively. However, the runoff response to temperature (temperature-induced changes and temperature elasticity) had the lowest predictability (<0.3). We inferred that the influence of many factors (climate, vegetation, soil, and land use/land cover) may assimilate RRCs, which leads to significant autocorrelation in space and thus high predictability, whereas RRCs with low predictability were independent of each other in space. Addor et al. (2018) considered that flow signatures with low predictability usually present a low global Moran's index in space, which represents autocorrelation with each other in space and tends to be a predictor of the predictability of flow signatures.

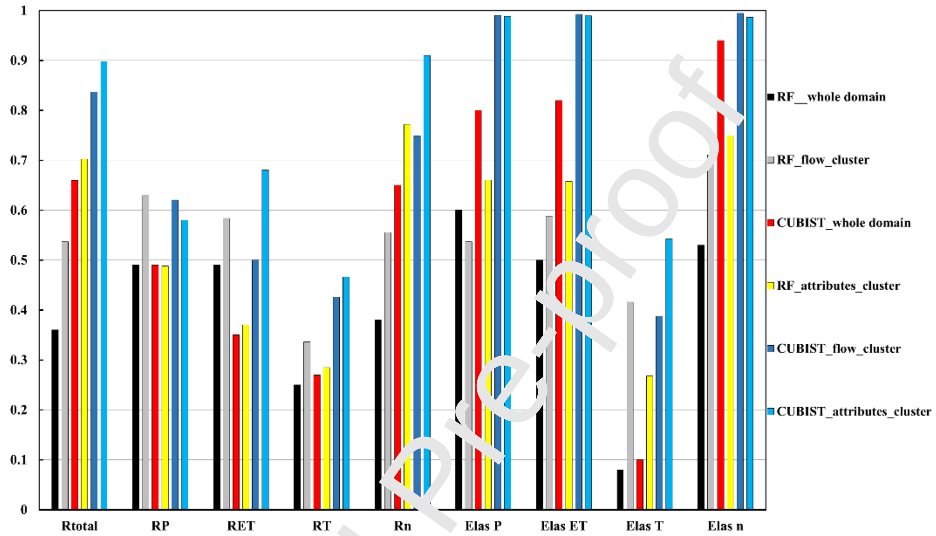


Figure 4. Performance comparison between random forest (RF) and CUBIST model for predicting runoff response changes over whole domain and within CA and FS based clusters. Performance of model within cluster for each of RRC is evaluated by the highest R^2 among all clusters. RF_whole_domain, RF_flow_cluster, RF_attributes_cluster refers to random model performance over whole domain, within FS-based and CA-based clusters, respectively. Similarly, CUBIST_whole_domain, CUBIST_flow_cluster, and CUBIST_attributes_cluster, represent CUBIST model performance.

Ten models were built for each RRC, yielding a total of 900 models for nine RRC components for each ML model. We chose the model with the best performance (i.e., the highest R^2) in all clusters for each RRC (Figures 4-5). It was observed that the classification of either CA-based or FS-based models can substantially improve model performance compared to the general model built in the entire domain, with a mean improvement of 39% in model performance (Figures 4-5). The CUBIST model presented better predictive skills than the random forest model within clusters. For example, for the random forest model, the highest R^2 in the CA-based classes was 0.71. While 0.89 for the CUBIST model for total runoff changes. We must acknowledge that either the random forest or CUBIST model has few differences between CA and FS based clusters (only the highest performance has been highlighted), which is expected because different clusters have different dominant features, and varied RRCs tend to be controlled by different factors that determine the predictability of the RRC. This may be the reason why there is no obvious rule in the model performance for the CA-based and FS-based clusters.

We chose the CUBIST model to analyze the difference in the predictability of RRC between clusters because of its better predictive skill. We manually divided model performance into four levels based on R^2: “extreme” [>0.8], “high” [0.6,0.8], “moderate” [0.4,0.6], and “low” [<0.4]. As shown in Figure 5, the predictability of the RRC between classes, either CA-based- or FS-based, exhibited substantial differences. For example, model performance in eight classes (except for class 5 and 7) CA-based was “Extreme” (R^2>0.8) for catchment characteristic elasticity, while there was no class within which model performance belonging to “High” (R^2>0.6) for precipitation, temperature-induced changes and temperature elasticity. FS-based clusters exhibited a slightly better performance than CA-based clusters, which is expected because this classification was based directly on flow behavior.

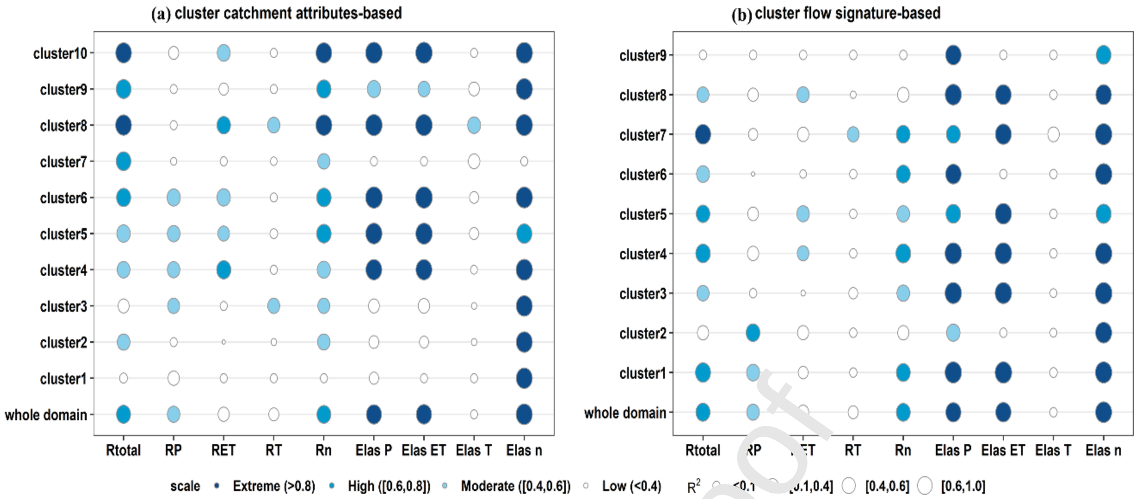


Figure 5. Performance of CUBIST model for predicting runoff response changes within each class and over the entire domain, catchments classified based on (a) CA, and (b) FS. Model performance is evaluated by R2. Rtotal, RP, RET, RT, Rn represents total changes of runoff, precipitation, potential evaporation, temperature, catchment characteristic-induced changes of runoff, respectively. Elas P, Elas ET, Elas T, and Elas n is the runoff sensitivity to the corresponding factors. Deeper the color, better the model performance.

3.3. Dominant controls on runoff response changes in the whole dataset and within clusters

The CUBIST model enables the evaluation of the importance of factors to RRC through attribute usage (importance), that is, the percentage of models where each RRC was used. We excluded the attributes where usage was less than 30% in the model over the whole domain (importance< 30%); the ranked importance of the attributes is shown in Figure 6. This analysis provided two insights: the difference in scores between a single attribute and the difference between classes. FS is the most important factor for the RRC over whole domain, excluding temperature-induced changes, with the rate of the flow duration curve, annual runoff, and runoff ratio being the first three important factors. The total changes were dominated by FS, topography, and soil characteristics, the first seven attributes were all related to them. The same is true for catchment-characteristic-induced changes. Climate plays an important role in the remains, but is less important for total changes and catchment characteristic-induced changes. However, temperature-induced changes were dominated by low precipitation duration. It should be noted that gauge longitude is a non-negligible factor for RRC and may be a cue for the spatial pattern of RRC over the CONUS (Patterson et al. 2013).

Influences on runoff sensitivity can be classified into two types: FS and climate, with the runoff ratio and annual runoff for FSs and the energy gradient (actual evaporation) for climate. Previous studies have shown that runoff is more sensitive to precipitation in arid regions (Sankarasubramanian et al. 2001, Fu et al. 2007). This finding is consistent with our results.

To understand the controls on RRC within the clusters, we selected the RRC whose model performance at least belongs to “High” level (R^2 >0.6) as the accuracy of importance is directly determined by the model performance. Resulting the first five attributes were selected based on their scores within each cluster. Considering that there is no substantial discrepancy in model performance between CA-based and FS-based classifications, for simplicity, we selected the CA-based classification (FS-based important features within clusters as seen in Figure S7 in Supporting Information S1). This procedure resulted in R^2 >0.6 in clusters 6,7,8,9 and10 for total changes, 4 and 8 for potential evaporation-induced changes, 5,6,8,9 and 10 for catchment characteristic-induced changes, 4,5,6,8 and10 for elasticity of precipitation and potential evaporation, and lastly 1,2,3,4,5,6,8,9 and10 for catchment characteristic elasticity; yet no cluster for precipitation, temperature-induced changes, and temperature elasticity (Figure7). We summarized the typical features for each RRC within the corresponding clusters in Table 2; where RRCs with R^2 <0.6 in clusters was not shown. On an average, the runoff ratio, annual runoff, 95th percentile runoff, and evaporation were the dominant factors for RRC through the catchment classification method. However, factors such as hydrology, soil, and topography should not be neglected either. In addition, there are some differences in the order of importance between CA-based and FS-based clusters as well as over the entire domain (Figure 6 and Figure 7), which may be caused by differences in model performance between them.

Table 2 Dominant attributes for RRC that satisfied selection criteria within clusters where model performance at least belongs to “High” level (R^2>0.6). Dominant attributes here refer to the catchment attributes and flow signatures that have the highest importance to corresponding RRC. Number in each column is the cluster, “R total”, “RET”, “Rn”, “Elas.P”, “Elas.ET”, “Elas.n” is total changes, potential evaporation-induced changes, catchment characteristic-induced changes, precipitation elasticity, potential evaporation elasticity, and catchment characteristic elasticity, respectively. ‘-’indicates that model performance within the corresponding cluster is less than 0.6, and not shown.

RRC	1	2	3	4	5	6	7	8	9	10

variables										
R total	-	-	-	-	-	95th percent runoff	95th percent runoff	evaporation	evaporation	Runoff ratio
RET	-	-	-	Annual runoff	-	-	-	Runoff ratio	-	-
Rn	-	-	-	-	95th percent runoff	95th percent runoff	-	evaporation	evaporation	Runoff ratio
Elas.P	-	-	-	Runoff ratio	Runoff ratio	Runoff ratio	-	Runoff ratio	-	Runoff ratio
Elas.ET	-	-	-	Runoff ratio	Runoff ratio	Runoff ratio	-	Runoff ratio	-	Runoff ratio
Elas.n	Runoff ratio	Runoff ratio	Annual runoff	Runoff ratio	Runoff ratio	evaporation	-	Annual runoff	Runoff ratio	Annual runoff

It should also be noted that climate tended to have a secondary effect on RRC over the entire dataset. This has been illustrated by Jehn et al. (2020), who linked this to the climate across catchments with high overlapped with other catchment attributes, such as vegetation and human intensification, thus resulting in a less clear role of climate on the RRC over the whole domain. In fact, the co-interaction between climate and landscape attributes produces a complex ecosystem in which a series of runoff generation processes exist (Gupta et al. 2014, Yin et al. 2017). Climate is the most important hydrological process in large-scale basins (Gupta et al. 2014), whereas geology and soil are important for small-scale basins (Troch et al. 2013, Li et al. 2018).

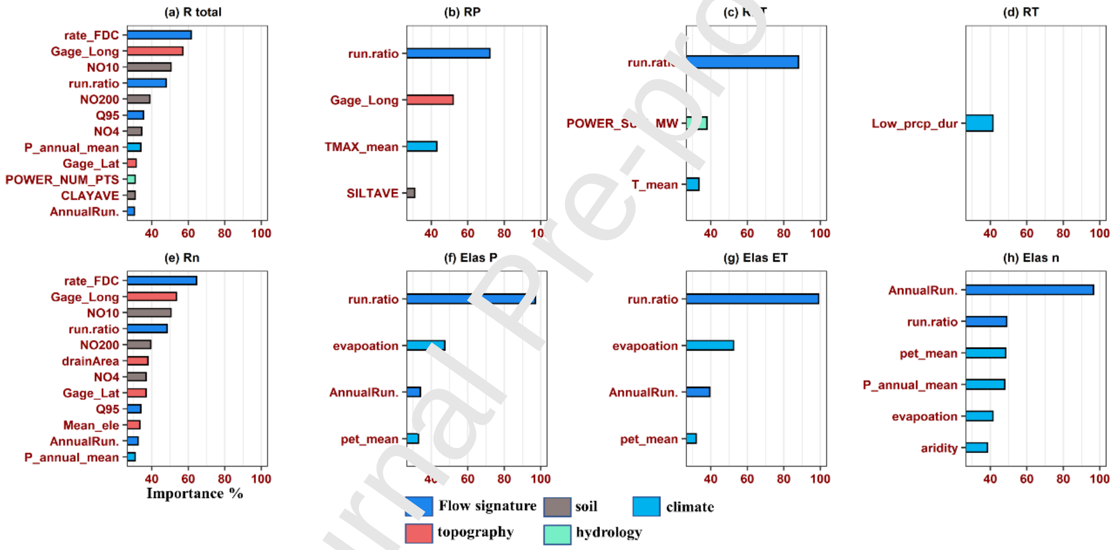


Figure 6. Importance of features evaluated by CUBIST model across whole domain, features are colored by catchment attribute and flow signature types.

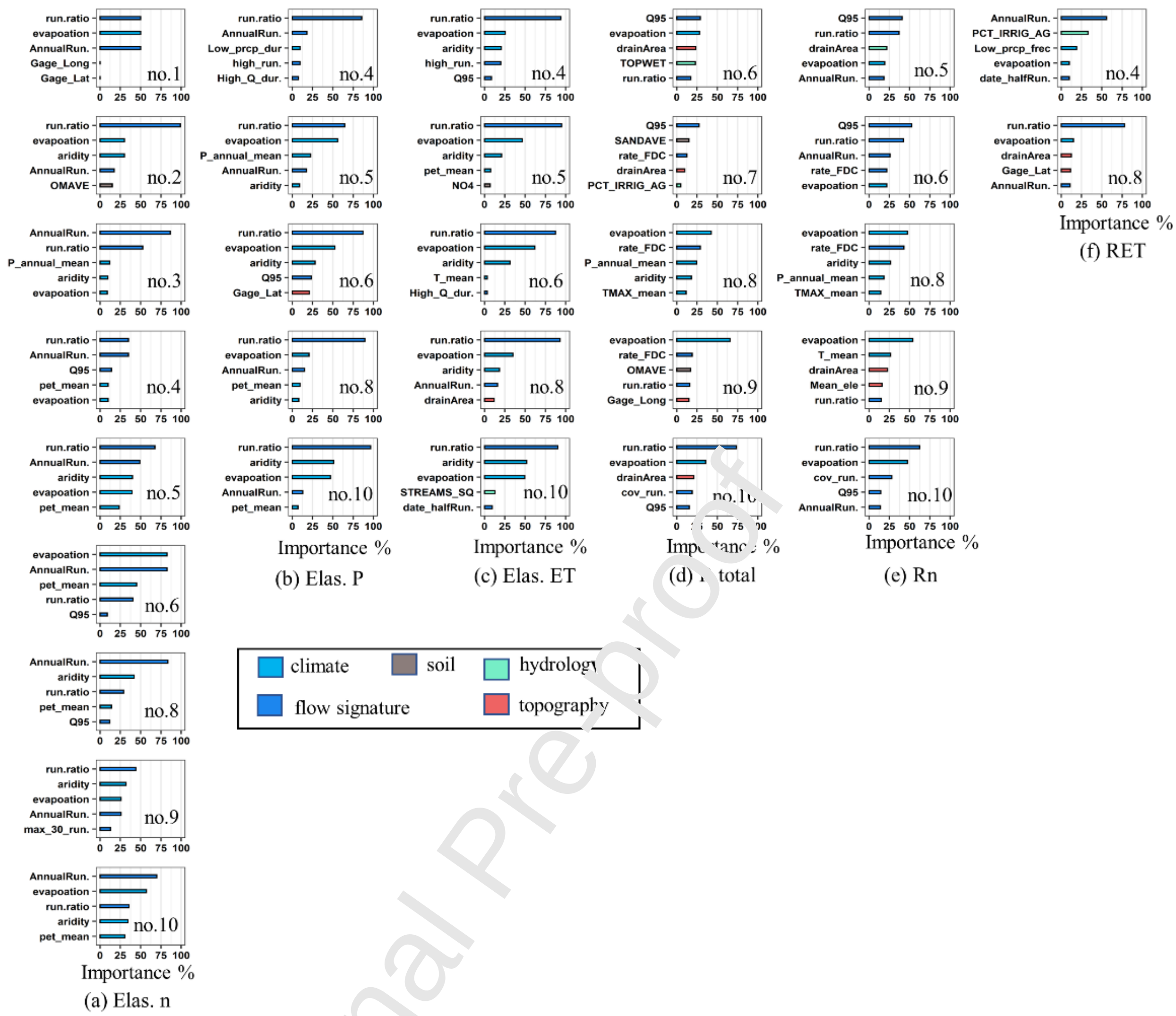


Figure 7. Importance of features evaluated by CUBIST model within CA-based clusters, features are colored by catchment attribute and flow signature types. Clusters in the figure have high or extreme R^2 (>0.6). Results of FS-based clusters are shown in Figure S7 in Supporting Information S1.

3.4. Runoff response changes pattern in dominant factors space

We explored the patterns of RRC in the space of dominant factors, as in the work of Knoben et al. (2018). Dominant factors in this study are defined as the top two important factors for RRC (with R^2>0.4 across the whole domain. Therefore, the RRC in terms of temperature-induced changes and temperature elasticity was excluded because their R^2 <0.4. As shown in Figure 8, the rate of flow duration curve and gauge longitude are the dominant factors for total changes, which are mainly located in west of the CONUS, while total changes tend to be independent of the rate of flow duration curve (Figure 8a), which is consistent with the findings of Gong et al. (2022). The sum of the precipitation and potential evaporation-induced changes are climate-induced runoff changes (Arrigoni et al. 2010, Ahn and Merwade 2014). As shown in Figure 8 b and c, the impact of the geographical gradient on climate-induced changes was clear, and high positive changes tended to be in the east, which was also explored by Arrigoni et al. (2010). The patterns of climate-induced changes tended to be independent of the runoff ratio. The role of human activities (construction of power sites) on potential evaporation-induced changes is worth noticing (Kotz et al. 2021). The pattern of catchment characteristic-induced changes is consistent with the total changes, which is reasonable because the shape parameter in the Budyko framework reflects the combined effects of landscapes and climate (Wu et al. 2017).

The runoff sensitivity exhibited a clear pattern across the dominant factors space. Climate elasticity (precipitation and potential evaporation elasticity) tends to be high in spaces of high evaporation and low runoff ratio and low in spaces of low evaporation and high runoff ratio. The pattern of catchment characteristic elasticity was consistent with climate elasticity, and the dominant factor space was the annual runoff and runoff ratio. Similar patterns of runoff sensitivity were expected because their dominant factors were correlated with each other. For example, the correlation between evaporation and runoff ratio was -0.6, and 0.7 for annual runoff and runoff ratio (Figure

S8 in Supporting Information S1).

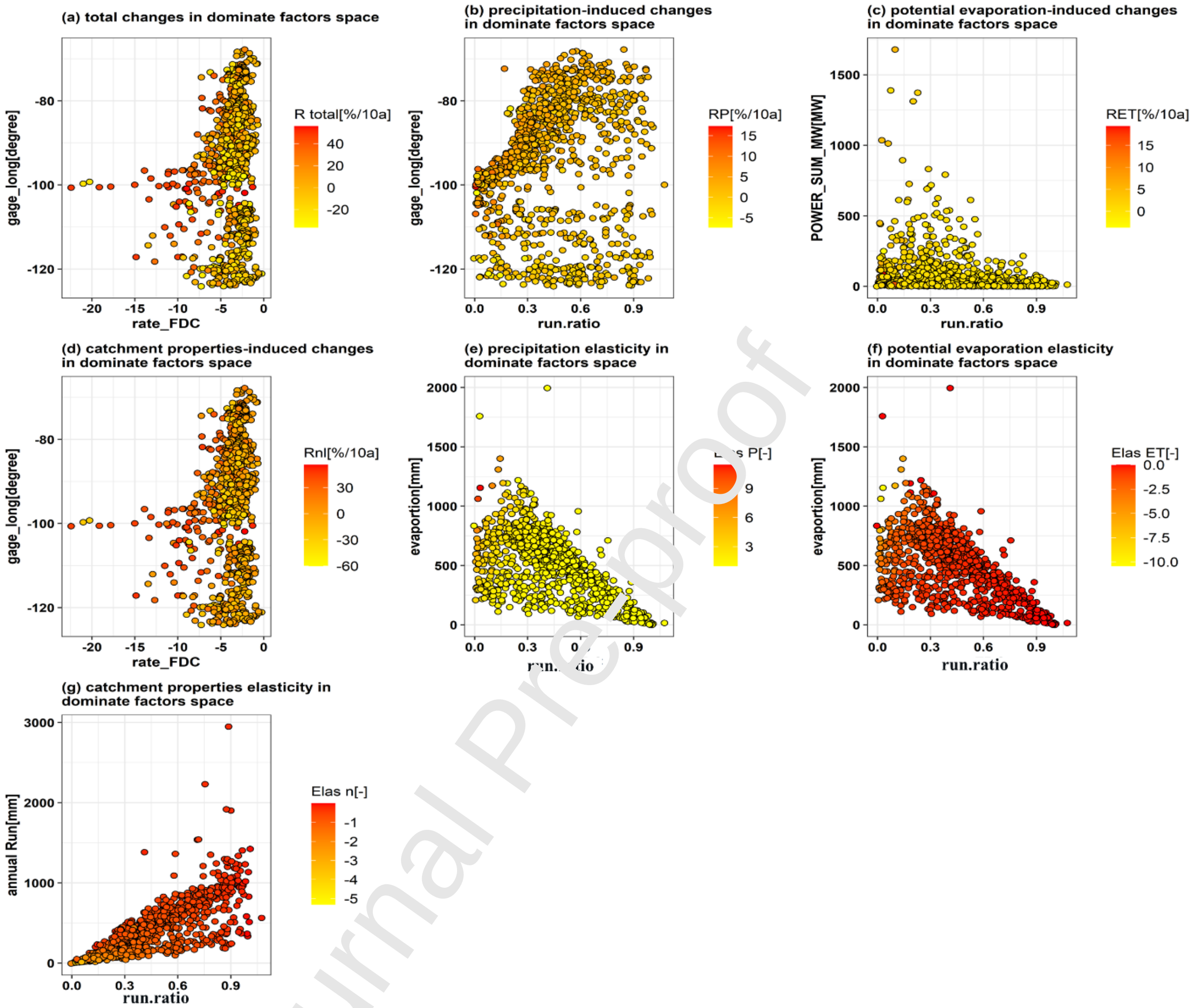


Figure 8. Runoff response changes in dominant factors space. (a) total changes (R total), (b) precipitation-induced changes (RP), (c) potential evaporation-induced changes (RET), (d) catchment attributes-induced changes (Rn), (e) precipitation elasticity (Elas P), (f) potential evaporation elasticity (Elas ET), (g) catchment properties elasticity (Elas n). Dominant factors have been defined by importance of factors evaluated by CUBIST model over whole domain (Figure 6).

4. Discussion

4.1. Uncertainty in the number of catchments between clusters

Catchments are synthetic systems in which the interaction between climatic and landscape processes produces a co-evolution of hydrological processes (Coopersmith et al., 2012;). Therefore, it is difficult to identify the dominant cause of alterations in hydrological response (Trancoso et al., 2017). Catchment classification may be a useful tool for distinguishing the dominant cause from the combined impact of various factors on hydrological behavior. However, there is large uncertainty in the number of catchments between clusters (Coopersmith et al. 2012, Kuentz et al. 2017, Jehn et al. 2020), as shown in Table 1. In the present study, for example, there were 175 catchments in Cluster 2 that were FS-based, whereas only one catchment was in Cluster 10. We attributed this difference to two factors: the uncertainty of large-scale databases, which may sometimes even be misinformative at high resolutions (Donnelly et al., 2012; Kauffeldt et al., 2013; Kuentz et al.,2017). This uncertainty may further obscure the classification of catchments over large-scale domains, and thus present ambiguous hydrological behavior. In addition, some strong externals, such as human alteration, that substantially affect hydrological variability have usually been ignored owing to their inability to access or difficult quantification. Hence, there is a need to investigate the impact of human activities on hydrological behavior. The catchments in the present study were mainly clustered in mountains, droughts, and plains (Table 1), similar to the results of Kuentz et al. (2017) and Jehn et al. (2020), where complex influences from climate, topography, geology, land use/land cover, and vegetation overlapped. One class has two or more features, instead of one important feature that controls the hydrological behavior over the CONUS (Table 1), which may be another possible cause of this indistinct behavior. For example, drought regions are usually represented by high energy and little rainfall, while dominant features for mountains and/or plains are not easily identified as several mixed influences, such as snowpacks, climate seasonality, land use change, and groundwater interaction govern it (Trancoso et al., 2017; Yang et al.,2022). Therefore, it is somewhat difficult for hydrologists to build homogenous groups of catchments with a single dominant attribute to explore hydrological behavior because, apart from overlapping influences, catchments would also self-organize due to the co-evolution of climate, vegetation, and landscape attributes (Burn, 1997).

To our knowledge, the impact of this difference in number of catchments between clusters on capturing of hydrological behavior is still ambiguous. To investigate the influence, we here used fuzzy clustering method to try to archive the aim, which maximizes the similarity between objects divided into the same cluster and minimizes the similarity between different clusters, in other words, the degree of membership is used to indicate the degree to which a sample belongs to a certain category.

Clusters from fuzzy clustering approach tend to be scattered over CONUS compared to agglomerative hierarchical algorithm in present study (see Fig. 2 and Figure S9 in Supporting Information S1), the patterns might be not following the ecoregions across CONUS. For example, both cluster 5 CA-based and 1 FS-based are scattered across different ecoregions (Figure S9 in Supporting Information S1). Which is different from previous classification results over CONUS. Jehn et al. (2020) found the clusters can follow the ecoregions well over CONUS, who used our method to divide 671 catchments into 10 classes. It is noted that fuzzy clustering also produced uncertainty in number of catchments between clusters, and this uncertainty is larger than when using agglomerative hierarchical method (see Table A1 where four clusters CA-based have <35 catchments, which are large difference compared to other clusters with >100 catchments, while two clusters CA-based <50 catchments for agglomerative hierarchical approach in Table 1). Therefore, we can infer that this uncertainty might be a cause of poor ability in capturing variety of hydrological behavior over CONUS, as following ecoregions is a signal of a good classification (Jehn et al., 2020). In addition, this uncertainty may also induce less connection between clusters and average catchment attributes. Some clusters can illustrate well mean catchment attributes, while others cannot. For example, class 1 CA-based has clear connection to low precipitation duration compared to others; while all clusters of ten have poor connection to climate seasonality (Figures S1-S4 in Supporting Information S1).

Even though uncertainty in numbers of catchments for clusters, typical features in each class tend to be robust. Most clusters CA/FS-based from either our method or fuzzy clustering have same typical features (Tables 1 and A1, details for catchment attributes and flow signatures in each class see Figures S1-S4 in Supporting Information S1). For example, the typical features of classes CA-based are related to soil characteristic, while mean half flow-date for classes FS-based. Therefore, it is not difficult to understand why there is same predictability of RRC between classes using agglomerative hierarchical and fuzzy algorithm (Fig.4 and Figure S5 in Supporting

Information S1).

Previous studies have pointed that fuzzy algorithm may solve the uncertainty in number of catchments between classes (Jehn et al. 2020). We here demonstrated that, however, fuzzy clustering caused same typical features and predictability of RRC as our method, but numbers of catchment within clusters still, and classification from fuzzy clustering method tend to be poor. However, no matter classification based on either agglomerative hierarchical or fuzzy method, the results can archive our needs of capturing dominant hydrological behavior, as the aim of hydrological classification is to distinguish hydrological behavior between clusters.

4.2. Understanding of runoff response changes with poor predictability

The response of runoff to temperature had the poorest predictability in the present study, but the results were robust in our large-sample study. Andréassian et al. (2016) demonstrated that the physical reasons for spatial patterns in runoff sensitivity are difficult to identify. We believe that a better understanding of the poor predictability of RRCs would improve the reliability of the predictions and be useful for assessing whether hydrological models can correctly capture streamflow behavior or not(Vano et al., 2015).

Sankarasubramanian et al. (2001) linked spatial variations in precipitation elasticity to the climate gradient, which was also reported by Fenta et al. (2017). Andréassian et al. (2016) inferred that the absence or presence of large groundwater aquifers can impact the spatial patterns of runoff response; however, more detailed comparative studies are still needed to explore the physical mechanism. In this study, we linked the runoff response to catchment attributes and flow signatures, but ambiguity still prevailed for some RRCs (e.g., temperature elasticity and temperature-induced changes). For those RRCs, neither the random forest nor the CUBIST model could capture their dominant factors. One explanation is that the non-parametric method for computing the response of runoff to temperature results in outliers when there is a very limited span of climatic factor anomalies (Andréassian et al., 2016), even though it is robust compared with other methods (Sankarasubramanian et al. 2001, Yang et al. 2008). However, this is inevitable in our large-sample study, and either process-based methods (e.g., SIMHYD, ABCD, and AWBM models) or empirical methods (e.g., non-parametric and regression methods) have shortcomings in computing the runoff response and resulting outliers (Andréassian et al., 2016).

However, this does not imply that poorly predictable RRCs should be neglected. In contrast, we suggest that they should be used carefully to capture the runoff response. Addor et al. (2018) found that low-flow metrics have the poorest predictability in space using random forest, but low flow is important for the ecology and simulation of rainfall-runoff relationships. In addition, the co-variation of different climatic factors is a problem in empirical (data-based) elasticity assessments. This makes it difficult to attribute streamflow changes to a single cause (Fu et al. 2007). Therefore, a better understanding of the driving processes is required.

5. Conclusions

This study explored the RRC behavior in terms of predictability and control over the entire dataset and within clusters through use method of catchment classification and

ML. RRC refers to the total changes, precipitation, potential evaporation, temperature, and catchment characteristic-induced changes in runoff, as well as the sensitivity of runoff to precipitation, potential evaporation, temperature, and catchment characteristics. Hydrological classification was used to group 1003 catchments into 10 clusters with similar catchment/flow behavior based on 56 CAs in terms of climate, topography, geology, hydrology, and soil properties, and 16 FSs with good predictability in space. Two MLs, random forest and the CUBIST model, were developed and compared in terms of model performance within clusters and over the whole domain. The optimal model was chosen to investigate the dominant controls of RRC within the clusters and over the entire dataset. Finally, we conclude the following:

1. Catchment classification can substantially improve the understanding of catchment behavior. 1003 catchments can be grouped into 10 clusters with similar flow and/or catchment behavior across the CONUS. Clusters based on either CA or FS partly follow the ecological regions in the U.S.. However, catchments with similar CA/FS behaviors can be quite distant from each other. The patterns of the CA-based and FS-based clusters in the climate index space are similar. Differences in the signals of forcing climate in catchment behavior may explain why catchments tend to show surprisingly similar behavior across many different climatic and landscape properties and why the most hydrologically similar catchment can be hundreds of kilometers away.
2. The CUBIST model was superior to the random forest model, either within or over the entire domain, and catchment classification substantially improved the

predictability of RRC across the CONUS (with a mean improvement of 39% in model performance, depending on clusters). On an average, the CUBIST model, with a mean R^2 of 0.56 is better than the random forest model with 0.41for predicting RRC over the whole domain, while 0.78 for CUBIST vs. 0.55 for random forest within clusters. The predictability of the RRC between clusters, either CA-based- or FS-based, exhibits relative differences. Runoff sensitivity can be predicted better than changes in runoff.

3. The flow signatures and climate were the dominant factors for RRC. There are some differences in the order of importance between the CA/FS-based clusters and the entire domain. On an average, the runoff ratio, annual runoff, and evaporation were the dominant factors for the RRC. The role of climate tended to be significant within a cluster, whereas it appeared to overlap over the entire domain. This is because the model performance was better within clusters than across the entire domain. Second, the signal of climatic forcing is superimposed more by other catchment attributes.
4. The RRC patterns in the dominant factor space are substantial. The patterns of runoff sensitivity were more obvious than those of runoff changes. High total changes/catchment characteristic-induced changes were mainly observed at 100°west longitude. Highly positive climate-induced changes tended to occur in the east. Climate/catchment characteristic elasticity tended to be high in spaces of high evaporation and low runoff ratios and low in spaces of low evaporation and high runoff ratios.

We found that clusters derived from both flow signatures and catchment attributes partly follow the ecoregions over the CONUS, and catchments within each cluster have approximately similar river flow regimes; thus climate or flow signatures are the dominant factors for the hydrological behavior of most catchments. This is of significance for water resource management and ecosystem protection, as policymaking and detailed actions generally need to be pointed out at the critical problem of the aim, which is the main contribution of our study. Ecosystems across regions are affected not only by landscape features, but also by runoff behaviors, such as root zone water storage, which determines the water uptake of plants, and baseflow, which restricts water quantity for humans and protects the water demand of aquatic organisms. We also found that the CUBIST model had better skills than the random forest model in predicting the RRC. CUBIST predicts response variables using non-linear relationships, and a series of segmented linear models in each leaf node of the model tree are combined to solve nonlinear problems effectively. However, random forest makes predictions using the mean of all training results at the final leaf node, which leads to bias in the new test set. The performance of CUBIST can be a reference for exploring regional variations in non-linear and non-monotonic controls on hydrological behavior and model structure, and thus can predict future hydrological regimes more accurately. Moreover, the methods and conclusions from this large-scale sample study could serve as a guide for investigating climate controls on hydrological regimes over Mediterranean regions, where climate seasonality is dominant, as well as for roles of geology and/or soil characteristics over small-scale catchments in other parts of the world.

We acknowledge that our results are dependent on the uncertainty of the classification of catchments, catchment attributes, the flow signatures considered, and the machine learning model used. Although the GAGES-II dataset provided excellent descriptions of different climatic and topographical characteristics for different catchment types, flow signatures with high spatial predictability resulted in meaningful clusters that could describe multiple aspects of the flow regime over different time scales. However, comprehensive datasets such as GAGES-II are still weak in obtaining conclusive clusters of catchments. In addition, methods such as fuzzy clustering cannot reduce uncertainty in the number of catchments between clusters. Therefore, for future research, we recommend examining a single cluster in greater depth and investigating the interaction between climate, topography, and soil over a single catchment, especially co-evolution. In addition, it is necessary to embed physical principles in machine learning models, especially when applying the model to ungauged basins, because the black box is weak in interpreting relationships between input and output, and requires a large amount of data to calculate and test the parameters.

Acknowledgments

This work was financially supported by the National Natural Science Foundation of China (U2243203, 51979069); the Fundamental Research Funds for the Central Universities (B200204029); National Natural Science Foundation of Jiangsu Province, China (BK20211202); and Research Council of Norway (FRINATEK Project 274310).

Appendix A

Table A1 Typical features of catchments classification. Catchments are classified by fuzzy cluster algorithm based on catchments attributes (CA) and flow signatures (FS) respectively. Typical feature refers to the signature and/or attribute of the cluster with the lowest coefficient of variation. Main regions of each cluster have not being highlighted as clusters did not follow ecoregions.

Cluster	Number for CA classification	Number for FS classification	Typical feature for CA classification	Typical feature for FS classification
1	222	152	Low bulk density	Later mean half flow-date
2	32	70	Low bulk density	High 95th percentile flow
3	102	5	High precipitation duration	Low flow
4	69	164	Low bulk density	Low 95th percentile flow
5	157	106	Low bulk density	High 95th percentile flow
6	299	161	High precipitation duration	Low 95th percentile flow
7	59	168	Low bulk density	Later mean half flow-date
8	18	35	Low bulk density	High 95th percentile flow
9	21	12	High soil thickness	High runoff ratio
10	24	130	Low bulk density	Later mean half flow-date

References

Atwood,A.,Hille, M.,Clark, M. K.,Rengers, F.,Ntarlagiannis, D.,Townsend, K. & West, A. J.(2023).Importance of subsurface water for hydrological response during storms in a post-wildfire bedrock landscape. Nature Communications,14(1): 3814-3814. doi:10.1038/S41467-023-39095-Z.

Addor, N., Nearing, G., Prieto, C., Newman, A. J., Le Vine, N., & Clark, M. P. (2018). A Ranking of Hydrological Signatures Based on Their Predictability in Space. Water Resources Research, 54(11), 8792-8812. https://doi.org/10.1029/2018wr022606.

Addor, N., Newman, A. J., Mizukami, N., & Clark, M. P. (2017a). The CAMELS data set: catchment attributes and meteorology for large-sample studies. Hydrology and Earth System Sciences, 21(10), 5293-5313. https://doi.org/10.5194/hess-21-5293-2017.

Addor, N., Newman, A. J., Mizukami, N., & Clark, M. P. (2017b). Catchment attributes for large-sample studies. Hydrology and Earth System Sciences, 21, 5293–5313. https://doi.org/10.5065/D6G73C3Q.

Ahmad H. N., Ashish S., Lucy M., Ian C. (2023). Modelling climate change impacts on the Brahmaputra streamflow resulting from changes in snowpack attributes. Journal of hydrology,603(Part B),126998. https://doi.org/10.1016/j.jhydrol.2021.126998.

Ahn, K. H., & Merwade, V. (2014). Quantifying the relative impact of climate and human activities on streamflow. Journal of Hydrology, 515, 257-266. https://doi.org/10.1016/j.jhydrol.2014.04.062.

Alvarez-Garreton, C., Mendoza, P. A., Boisier, J. P., Addor, N., Galleguillos, M., Zambrano-Bigiarini, M., Lara, A., Puelma, C., Cortes, G., Garreaud, R., McPhee, J., & Ayala, A. (2018). The CAMELS-CL dataset: catchment attributes and meteorology for large sample studies – Chile dataset. Hydrology and Earth System Sciences, 22(11), 5817-5846. https://doi.org/10.5194/hess-22-5817-2018.

Andréassian, V., Coron, L., Lerat, J., & Le Moine, N. (2016). Climate elasticity of streamflow revisited – an elasticity index based on long-term hydrometeorological records. Hydrology and Earth System Sciences, 20(11), 4503-4524. https://doi.org/10.5194/hess-20-4503-2016.

Arrigoni, A. S., Greenwood, M. C., & Moore, J. N. (2010). Relative impact of anthropogenic modifications versus climate change on the natural flow regimes of rivers in the Northern Rocky Mountains, United States. Water Resources Research, 46(12). https://doi.org/10.1029/2010wr009162.

Auch, R., Drummond, M., Xian, G., Sayler, K., Acevedo, W., & Taylor, J. (2016). Regional Differences in Upland Forest to Developed (Urban) Land Cover Conversions in the Conterminous U.S., 1973–2011. Forests, 7(12). https://doi.org/10.3390/f7070132.

Bai, P., Liu, X., Zhang, D., & Liu, C. (2019). Estimation of the Budyko model parameter for small basins in China. Hydrological Processes, 34(1), 125-138. https://doi.org/10.1002/hyp.13577.

Bajirao, T.S., Kumar, P.(2022). Quantification of impact of spatio-temporal variability of land use/land cover on runoff generation using modified NRCS-CN method. Arab

J Geosci 15, 610. https://doi.org/10.1007/s12517-022-09931-5.

Balha, A., Singh, A., Pandey, S. et al.(2023). Assessing the Impact of Land-Use Dynamics to Predict the Changes in Hydrological Variables Using Effective Impervious Area (EIA). Water Resource Management, 37: 3999–4014. https://doi.org/10.1007/s11269-023-03536-7.

Basagaoglu, H., Chakraborty, D., Do Lago, C., Gutierrez, L., Sahinli, M. A., Giacomoni, M., Furl, C., Mirchi, A., Moriasi, D., & Sengor, S. S. (2022). A Review on Interpretable and Explainable Artificial Intelligence in Hydroclimatic Applications. Water, 14(8), Article 1230. https://doi.org/10.3390/w14081230.

Basri, H., Sufardi, S., Helmi, H., Syakur, S., Sugianto, S., Azmeri, A., Helmi, H. (2023). Drought and water availability analysis for irrigation and household water needs in the Krueng Jrue sub-watershed. PeerJ, 11,e14830. https://doi.org/10.7717/peerj.14830.

Beevers, L.; Bedinger, M., McClymont, K.; Visser-Quinn, A.(2021). Resilience in Complex Catchment Systems. Water, 13(4): 541. https://doi.org/10.3390/w13040541.

Bharat, S.& Mishra,V.(2021). Runoff sensitivity of Indian sub-continental river basins. Science of The Total Environment,766,142642. doi.org/10.1016/j.scitotenv.2020.142642.

Blöschl, G., Bierkens, M. F. P., Chambel, A., Cudennec, C., Destouni, G., Fiori, A., Kirchner, J. W., McDonnell, J. J., Savenije, H. H. G., Sivapalan, M., Stumpp, C., Toth, E., Volpi, E., Carr, G., Lupton, C., Salinas, J., Széles, B., Viglione, A., Aksoy, H., Zhang, Y. (2019). Twenty-three unsolved problems in hydrology (UPH) – a community perspective. Hydrological Sciences Journal, 64(10), 1141-1158. https://doi.org/10.1080/02626667.2019.1620507

Booker, D. J., & Woods, R. A. (2014). Comparing and combining physically-based and empirically-based approaches for estimating the hydrology of ungauged catchments. Journal of Hydrology, 508, 227-239. https://doi.org/10.1016/j.jhydrol.2013.11.007.

Breiman L. Classification and regression trees[M]. NewYork: Chapman and Hall, 1984.

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

Budyko, M. I. (1958), The Heat Balance ofthe Earth's Surface, U.S. Dep. ofCommer., Washington, D. C.

Budyko, M. I. (1974). Climate and life. San Diego, CA: Academic Press.

Carmona, A. M., Sivapalan, M., Yaeger, M. A., & Poveda, G. (2014). Regional patterns of interannual variability of catchment water balances across the continental U.S.: A Budyko framework. Water Resources Research, 50(12), 9177-9193. https://doi.org/10.1002/2014wr016013.

Burn, D. H.: Catchment similarity for regional flood frequency analysis using seasonality measures, J. Hydrol., 202, 212–230, https://doi.org/10.1016/S0022-1694(97)00068-1, 1997.

Chaney, N. W., Wood, E. F., McBratney, A. B., Hempel, J. W., Nauman, T. W., Brungard, C. W., & Odgers, N. P. (2016). POLARIS: A 30-meter probabilistic soil series map of the contiguous United States. Geoderma, 274, 54-67. https://doi.org/10.1016/j.geoderma.2016.03.025

Chang, F.J., Chang, L.C., Chen, J.F.(2023). Artificial Intelligence Techniques in Hydrology and Water Resources Management. Water, 15(10):1846. https://doi.org/10.3390/w15101846.

Cheng, S., Cheng, L., Qin, S., Zhang, L., Liu, P., Liu, L., Xu, Z., & Wang, Q. (2022). Improved Understanding of How Catchment Properties Control Hydrological Partitioning Through Machine Learning. Water Resources Research, 58(4). https://doi.org/10.1029/2021wr031412.

Cinkus, G., Mazzilli, N. & Jourde, H.(2023). KarstID: an R Shiny application for the analysis of karst spring discharge time series and the classification of karst system hydrological functioning. Environ Earth Sci, 82: 136. https://doi.org/10.1007/s12665-023-10830-5

Coopersmith, E., Yaeger, M. A., Ye, S., Cheng, L., & Sivapalan, M. (2012). Exploring the physical controls of regional patterns of flow duration curves – Part 3: A catchment classification system based on regime curve indicators. Hydrology and Earth System Sciences, 16(11), 4467-4482. https://doi.org/10.5194/hess-16-4467-2012.

Crawford, N. H., & Linsley, R. K. (1966). Digital simulation in hydrology: Stanford Watershed Model IV, Stanford University Department of Civil Engineering, Technical Report 39 (210 pp.).

Apley,D.W. Zhu,J.(2020). Visualizing the Effects of Predictor Variables in Black Box

Supervised Learning Models, Journal of the Royal Statistical Society Series B: Statistical Methodology, 82(4) , 1059–1086, https://doi.org/10.1111/rssb.12377.

Dennison, P. E., Brewer, S. C., Arnold, J. D., & Moritz, M. A. (2014). Large wildfire trends in the western United States, 1984-2011. Geophysical Research Letters, 41(8), 2928-2933. https://doi.org/10.1002/2014gl059576.

Di, Z., Duan, Q., Gong, W., Ye, A., & Miao, C. (2017). Parametric sensitivity analysis of precipitation and temperature based on multi-uncertainty quantification methods in the Weather Research and Forecasting model. Science China Earth Sciences, 60(5), 876-898. https://doi.org/10.1007/s11430-016-9021-6.

Donnelly, C, Andersson, J. C. M., and Arheimer, B.: Using flow signatures and catchment similarities to evaluate a multi-basin model (E-HYPE) across Europe, Hydr. Sci. J., 61, 255–273, 2016

Falcone, J. A. (2011). GAGES II: Geospatial attributes of gages for evaluating streamflow. Retrieved from https://water.usgs.gov/GIS/metadata/usgswrd/XML/gagesII_Sept2011.xml.

Fang, X. & Pomeroy J. W..(2023).Simulation of the impact of future changes in climate on the hydrology of Bow River headwater basins in the Canadian Rockies. Journal of Hydrology,620(PB): 129566. doi:10.1016/J.JHYDROL.2023.129566.

Fenta, A. A., Yasuda, H., Shimizu, K., & Haregeweyn, N. (2017). Response of streamflow to climate variability and changes in human activities in the semiarid highlands of northern Ethiopia. Regional Environmental Change, 17(4), 1229-1240. https://doi.org/10.1007/s10113-017-1103-y.

Fernández-Delgado,M., Sirsat, M.S.,Cernadas,E.,Alawadi,S.,Barro,S.&Febrero-Bande,M.(2018).An extensive experimental survey of regression methods. Neural Networks,111(2018),11-34. doi:10.1016/j.neunet.2018.12.010.

Floriancic, M. G., Meerveld, I., Smoorenburg, M., Margreth, M., Naef, F., Kirchner, J. W., & Molnar, P. (2018). Spatio-temporal variability in contributions to low flows in the high Alpine Poschiavino catchment. Hydrological Processes, 32(26), 3938-3953. https://doi.org/10.1002/hyp.13302.

Forbes, W. L., Mao, J., Ricciuto, D. M., Kao, S. C., Shi, X., Tavakoly, A. A., Jin, M., Guo, W., Zhao, T., Wang, Y., Thornton, P. E., & Hoffman, F. M. (2019). Streamflow in the Columbia River Basin: Quantifying Changes Over the Period 1951-2008 and Determining the Drivers of Those Changes. Water Resources Research, 55(8), 6640-6652. https://doi.org/10.1029/2018wr024256.

Fu, G., Charles, S. P., & Chiew, F. H. S. (2007). A two-parameter climate elasticity of streamflow index to assess climate change effects on annual streamflow. Water Resources Research, 43(11). https://doi.org/10.1029/2007wr005890.

Gao, K.,Yang, Y.,Gil, J. & Qu, X.(2023).Data-driven interpretation on interactive and nonlinear effects of the correlated built environment on shared mobility. Journal of Transport Geography,110, 103604. doi:10.1016/J.JTRANGEO.2023.103604.

Ghamariadyan, M., Imteaz, M.A. (2021). A wavelet artifcial neural network method for medium-term rainfall prediction in Queensland (Australia) and the Comparisons with conventional methods. Int J Climato,l41(S1):E1396–E1416.

Gong, X. L., Xu, A. Q., Du, S. P., & Zhou, Y. A. (2022). Spatiotemporal variations in the elasticity of runoff to climate change and catchment characteristics with multi-timescales across the contiguous United States. Journal of Water and Climate Change, 13(3), 1408-1424. https://doi.org/10.2166/wcc.2022.242.

Gupta, H. V., Perrin, C., Blöschl, G., Montanari, A., Kumar, R., Clark, M., & Andréassian, V. (2014). Large-sample hydrology: a need to balance depth with breadth. Hydrology and Earth System Sciences, 18(2), 463-477. https://doi.org/10.5194/hess-18-463-2014.

Haddad,K.& Rahman,A.(2023). Dimensionality reduction for regional flood frequency analysis: Linear versus nonlinear methods. Hydrological process, 37(4), 14864. https://doi.org/10.1002/hyp.14864.

J. F. A..(1914).CUBIST SCIENCE. Science(1022). doi:10.1126/science.40.1022.167.

Hartigan, J. A. and Wong, M. A.: Algorithm AS 136: A K-Means Clustering Algorithm, J. Roy. Stat. Soc. C-Appl., 28, 100–108, 1979.

He,D.,Jiang,J.,Ma,T.,Yang,G.,Westphal,C.,Garcia-Luna-Aceves,J.,Xia,S.(2021). CUBIST: High-Quality 360-Degree Video Streaming Services via Tile-based Edge Caching and FoV-Adaptive Prefetching, 2021IEEE International Conference on Web Services (ICWS), 208-218. doi: 10.1109/ICWS53863.2021.00039.

Gao,H., Birkel,C.,Hrachowitz,M., Tetzlaff,D., Soulsby,C. & Savenije, H.(2019).A simple topography-driven and calibration-free runoff generation module. Hydrology and Earth System Sciences,23(2): 787-809. doi:10.5194/hess-23-787-2019.

Jehn, F. U., Bestian, K., Breuer, L., Kraft, P., & Houska, T. (2020). Using hydrological and climatic catchment clusters to explore drivers of catchment behavior. Hydrology and Earth System Sciences, 24(3), 1081-1100. https://doi.org/10.5194/hess-24-1081-2020.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning with applications in R. New York: Springer.

Kam, J., & Sheffield, J. (2015). Changes in the low flow regime over the eastern United States (1962–2011): variability, trends, and attributions. Climatic Change, 135(3-4), 639-653. https://doi.org/10.1007/s10584-015-1574-0.

Kauffeldt, A., Halldin, S., Rodhe, A., Xu, C.-Y., and Westerberg, I. K.: Disinformative data in large-scale hydrological modelling, Hydrol. Earth Syst. Sci., 17, 2845–2857, https://doi.org/10.5194/hess-17-2845-2013, 2013.

Khaledian, Y., ,& Miller,B. (2020).Selecting appropriate machine learning methods for digital soil mapping. Applied Mathematical Modelling, 81(C),401-418. doi:10.1016/j.apm.2019.12.016

Knoben, W. J. M., Woods, R. A., & Freer, J. E. (2018). A Quantitative Hydrological Climate Classification Evaluated With Independent Streamflow Data. Water Resources Research, 54(7), 5088-5109. https://doi.org/10.1029/2018wr022913.

Kotz, M., Wenz, L., Stechemesser, A., Kalkuhl, M., & Levermann, A. (2021). Day-to-day temperature variability reduces economic growth. Nature Climate Change, 11(4), 319-325. https://doi.org/10.1038/s41558-020-00985-5.

Kuentz, A., Arheimer, B., Hundecha, Y., & Wagener, T. (2017). Understanding hydrologic variability across Europe through catchment classification. Hydrology and Earth System Sciences, 21(6), 2863-2879. https://doi.org/10.5194/hess-21-2863-

2017.

Li, Q., Wei, X., Yang, X., Giles-Hansen, K., Zhang, M., & Liu, W. (2018). Topography significantly influencing low flows in snow-dominated watersheds. Hydrology and Earth System Sciences, 22(3), 1947-1956. https://doi.org/10.5194/hess-22-1947-2018.

Li, Q. S., Gao, H. S., Chai, S., He, Z. Y., & Ji, G. X. (2022). Attribution analysis of different driving forces on vegetation and streamflow variation in the Jialing River Basin, China. Open Geosciences, 14(1), 1513-1523. https://doi.org/10.1515/geo-2022-0443.

Maftouh, A., El Fatni, O., El Hajjaji, S., Jawish, M. W., & Sillanpaa, M. (2023). Comparative Review of Different Adsorption Techniques Used in Heavy Metals Removal in Water. Biointerface Research in Applied Chemistry, 13(4), 19, Article 397. https://doi.org/10.33263/briac134.397

McCabe, G. J., & Wolock, D. M. (2014). Spatial and temporal patterns in conterminous United States streamflow characteristics. Geophysical Research Letters, 41(19), 6889-6897. https://doi.org/10.1002/2014gl061980.

Moratiel, R., Bravo, R., Saa, A., Tarquis, A. M., & Almorox, J. (2020). Estimation of evapotranspiration by the Food and Agricultural Organization of the United Nations (FAO) Penman–Monteith temperature (PMT) and Hargreaves–Samani (HS) models under temporal and spatial criteria – a case study in Duero basin (Spain). Natural Hazards and Earth System Sciences, 20(3), 859-875. https://doi.org/10.5194/nhess-20-859-2020.

O'Sullivan Cherie M.,Ghahramani Afshin,Deo Ravinesh C. & Pembleton Keith G.(2022).Pattern recognition describing spatio-temporal drivers of catchment classification for water quality. The Science of the total environment, 861,160240. doi:10.1016/J.SCITOTENV.2022.160240.

Padrón, R. S., Gudmundsson, L., Greve, P., & Seneviratne, S. I. (2017). Large-Scale Controls of the Surface Water Balance Over Land: Insights From a Systematic Review and Meta-Analysis. Water Resources Research, 53(11), 9659-9678. https://doi.org/10.1002/2017wr021215

Pal, M. (2005). Random forest classifier for remote sensing classification. International Journal of Remote Sensing, 26(1), 217-222. https://doi.org/10.1080/01431160412331269698

Patterson, L. A., Lutz, B., & Doyle, M. W. (2013). Climate and direct human contributions to changes in mean annual streamflow in the South Atlantic, USA. Water Resources Research, 49(11), 7278-7291. https://doi.org/10.1002/2013wr014618.

PRISM Climate Group (2016), Oregon State University [internet]. Descriptions of PRISM spatial climate datasets for the conterminous United States. [Available at http://www.prism.oregonstate.edu/documents/PRISM_datasets.pdf, (Cited 25 July 2016).]

Quinlan, J. R. (1992). Learning with continuous classes. In Paper presented at proceedings AI'92, 5th Australian conference on artificial intelligence.

Rice, J. S., Emanuel, R. E., & Vose, J. M. (2016). The influence of watershed characteristics on spatial patterns of trends in annual scale streamflow variability in the continental U.S. Journal of Hydrology, 540, 850-860. https://doi.org/10.1016/j.jhydrol.2016.07.006.

Sankarasubramanian, A., Vogel, R. M., & Limbrunner, J. F. (2001). Climate elasticity of streamflow in the United States. Water Resources Research, 37(6), 1771-1781. https://doi.org/10.1029/2000wr900330.

Senatore, A., Parrello, C., Almorox, J., & Mendicino, G. (2020). Exploring the Potential of Temperature-Based Methods for Regionalization of Daily Reference Evapotranspiration in Two Spanish Regions. Journal of Irrigation and Drainage Engineering, 146(3). https://doi.org/10.1061/(asce)ir.1943-4774.0001447.

Shahid, M., Rahman, K.U., Haider, S. et al. (2021). Quantitative assessment of regional land use and climate change impact on runoff across Gilgit watershed. Environ Earth Sci 80, 743. https://doi.org/10.1007/s12665-021-10032-x.

Shao, M., Zhao, G., Kao, S. C., Cuo, L., Rankin, C., & Gao, H. (2020). Quantifying the effects of urbanization on floods in a changing environment to promote water security — A case study of two adjacent basins in Texas. Journal of Hydrology, 589. https://doi.org/10.1016/j.jhydrol.2020.125154.

Shen, Q., Gao, G., Han, F., Xiao, F., Ma, Y., Wang, S., & Fu, B. (2018). Quantifying the effects of human activities and climate variability on vegetation cover change in a hyper-arid endorheic basin. Land Degradation & Development, 29(10), 3294-3304. https://doi.org/10.1002/ldr.3085.

Slater, L. J., Arnal, L., Boucher, M.-A., Chang, A. Y.-Y., Moulds, S., Murphy, C., Nearing, G., Shalev, G., Shen, C., Speight, L., Villarini, G., Wilby, R. L., Wood, A., and Zappa, M.(2023). Hybrid forecasting: blending climate predictions with AI models, Hydrol. Earth Syst. Sci., 27, 1865–1889, https://doi.org/10.5194/hess-27-1865-2023.

Zhang S.Y., Thian Y.G., Andrew B.G. B., Zhang G.X. (2023). Evaluation of the impact of climate change on the streamflow of major pan-Arctic river basins through machine learning models. Journal of hydrology, 619,129295. https://doi.org/10.1016/j.jhydrol.2023.129295.

Sivapalan, M. (2003). Prediction in ungauged basins: a grand challenge for theoretical hydrology. Hydrological Processes, 17(15), 3163-3170. https://doi.org/10.1002/hyp.5155

Stein, L., Clark, M. P., Knoben, W. J. M., Pianosi, F., & Woods, R. A. (2021). How Do Climate and Catchment Attributes Influence Flood Generating Processes? A Large-Sample Study for 671 Catchments Across the Contiguous USA. Water Resources Research, 57(4). https://doi.org/10.1029/2020wr028300.

Trancoso, R., Phinn, S., McVicar, T. R., Larsen, J. R., & McAlpine, C. A. (2017). Regional variation in streamflow drivers across a continental climatic gradient. Ecohydrology, 10(3). https://doi.org/10.1002/eco.1816.

Tim, A.,Ephraim, M.,Douglas, R., Hardy,D., Hemp, A., & Thomas, N.(2015).Evaluating machine learning approaches for the interpolation of monthly air temperature at Mt. Kilimanjaro, Tanzania. Spatial Statistics,14 (2015),91-113. doi:10.1016/j.spasta.2015.05.008.

Troch, P. A., Carrillo, G., Sivapalan, M., Wagener, T., & Sawicz, K. (2013). Climate-vegetation-soil interactions and long-term hydrologic partitioning: signatures of catchment co-evolution. Hydrology and Earth System Sciences, 17(6), 2209-2217. https://doi.org/10.5194/hess-17-2209-2013.

Vano, J. A., Nijssen, B., & Lettenmaier, D. P. (2015). Seasonal hydrologic responses to climate change in the Pacific Northwest. Water Resources Research, 51(4), 1959-1976. https://doi.org/10.1002/2014wr015909.

Viscarra Rossel, R. A., Lee, J., Behrens, T., Luo, Z., Baldock, J., & Richards, A. (2019). Continental-scale soil carbon composition and vulnerability modulated by regional environmental controls. Nature Geoscience, 12(7), 547–552. https://doi.org/10.1038/s41561-019-0373-z

Wagener, T., Sivapalan, M., Troch, P., and Woods, R.: Catchment Classification and Hydrologic Similarity. Geography Compass, 1, 901–931, https://doi.org/10.1111/j.1749-8198.2007.00039.x, 2007.

Wang, D., & Hejazi, M. (2011). Quantifying the relative contribution of the climate and direct human impacts on mean annual streamflow in the contiguous United States. Water Resources Research, 47(10). https://doi.org/10.1029/2010wr010283.

Wang, H. and K. He (2017). "Sensitivity Analysis of the Effects of Climate Change on Streamflow Using Climate Elasticity in the Luan River Basin, China." Polish Journal of Environmental Studies **26**(2): 837-845. https://doi.org/10.15244/pjoes/66715.

Wang, H. & Ni, W.(2023).Response and prediction of unsaturated permeability of loess to microstructure. Geomechanics and Geophysics for Geo-Energy and Geo-

Resources,9(1). doi:10.1007/S40948-023-00541-3.

Wang, J., Gao, Y., & Wang, S. (2018). Assessing the response of runoff to climate change and human activities for a typical basin in the Northern Taihang Mountain, China. Journal of Earth System Science, 127(3). https://doi.org/10.1007/s12040-018-0932-5.

Wang, M., Liu, M., Zhang, D., Qi, J., Fu, W., Zhang, Y., ... & Tan, S. K. (2023). Assessing and optimizing the hydrological performance of Grey-Green infrastructure systems in response to climate change and non-stationary time series. Water Research, 232, 119720.

Wang,S.,Liu, T., Qian, X.,Wang, H.,Wang, X.,Wei, S. & Chen, H.(2023).Contamination of the urban river network with perfluoroalkyl acids (PFAAs) introduced during river regulations. The Science of the total environment,893(2023): 164775-164775.doi:10.1016/J.SCITOTENV.2023.164775.

Ward Jr., J. H.: Hierarchical grouping to optimize an objective function, J. Am. Stat. Assoc., 58, 236–244, 1963.

Westerberg, I. K., & McMillan, H. K. (2015). Uncertainty in hydrological signatures. Hydrology and Earth System Sciences, 19(9), 3951-3968. https://doi.org/10.5194/hess-19-3951-2015.

Wu, J., Miao, C., Wang, Y., Duan, Q., & Zhang, X. (2017). Contribution analysis of the long-term changes in seasonal runoff on the Loess Plateau, China, using eight Budyko-based methods. Journal of Hydrology, 545, 263-275. https://doi.org/10.1016/j.jhydrol.2016.12.050.

Wu, J., Miao, C., Zhang, X., Yang, T., & Duan, Q. (2017). Detecting the quantitative hydrological response to changes in climate and human activities. Sci Total Environ, 586, 328-337. https://doi.org/10.1016/j.scitotenv.2017.02.010.

Yaeger, M., Coopersmith, E., Ye, S., Cheng, L., Viglione, A., and Sivapalan, M.: Exploring the physical controls of regional patterns of flow duration curves – Part 4: A synthesis of empirical analysis, process modeling and catchment classification, Hydrol. Earth Syst. Sci., 16, 4483–4498, https://doi.org/10.5194/hess-164483-2012, 2012.

Yang, H., Yang, D., Lei, Z., & Sun, F. (2008). New analytical derivation of the mean annual water-energy balance equation. Water Resources Research, 44(3). https://doi.org/10.1029/2007wr006135

Yang, K., Hou, H., Li, Y., Chen, Y., Wang, L., Wang, P., & Hu, T. (2022). Future urban waterlogging simulation based on LULC forecast model: A case study in Haining City, China. Sustainable Cities and Society, 87, 104167.

Yin, J., He, F., Xiong, Y. J., & Qiu, G. Y. (2017). Effects of land use/land cover and climate changes on surface runoff in a semi-humid and semi-arid transition zone in northwest China. Hydrology and Earth System Sciences, 21(1), 183-196. https://doi.org/10.5194/hess-21-183-2017.

Yin, H.,Zhu, W.,Zhang, X.,Xing, Y.,Xia, R.,Liu, J. & Z, Y.(2023).Runoff predictions in new-gauged basins using two transformer-based models. Journal of Hydrology,622(PA): 129684. doi:10.1016/J.JHYDROL.2023.129684.

Zhang, X., Zhang, L., Zhao, J., Rustomji, P., & Hairsine, P. (2008). Responses of streamflow to changes in climate and land use/cover in the Loess Plateau, China. Water Resources Research, 44(7). https://doi.org/10.1029/2007wr006711.

Zhang, Y., Ye, A., You, J., & Jing, X. (2021). Quantification of human and climate contributions to multi-dimensional hydrological alterations: A case study in the Upper Minjiang River, China. Journal of Geographical Sciences, 31(8), 1102-1122. https://doi.org/10.1007/s11442-021-1887-z.

Zhao, G., Pang, B., Xu, Z., Peng, D., & Xu, L. (2019). Assessment of urban flood susceptibility using semi-supervised machine learning model. Science of the Total Environment, 659, 940-949.

Zutta, B. R.,Salinas, N.,Cosio, E. G.,Tito, R., Aragón, S.,Nina-Quispe, A. & Roman-Cuesta, R. M.(2023).Satellite-derived forest canopy greenness shows differential drought vulnerability of secondary forests compared to primary forests in Peru. Environmental Research Letters,18(6), 1748-9326. doi:10.1088/1748-9326/ACC8EA.

1. Catchments are grouped into 10 clusters with substantial features over ecoregions;
2. Data-driven model's performance in RRC can be improved through catchment classification;
3. Flow signatures are the most important for RRC followed by climate forcing;
4. Uncertainty in number of catchments between clusters need to be considered