Hydrology and Earth System Sciences

Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA

K. Sawicz[1], T. Wagener[1], M. Sivapalan[2,3], P. A. Troch[4], and G. Carrillo[4]

[1]Department of Civil and Environmental Engineering, The Pennsylvania State University, 212 Sackett Building,
University Park, PA16802, USA
[2]Department of Civil and Environmental Engineering and Department of Geography,
University of Illinois-Urbana Champaign, USA
[3]Department of Water Management, Delft University of Technology, The Netherlands
[4]Department of Hydrology and Water Resources, The University of Arizona, Tucson, AZ, USA

Abstract. Hydrologic similarity between catchments, derived from similarity in how catchments respond to precipitation input, is the basis for catchment classification, for transferability of information, for generalization of our hydrologic understanding and also for understanding the potential impacts of environmental change. An important question in this context is, how far can widely available hydrologic information (precipitation-temperature-streamflow data and generally available physical descriptors) be used to create a first order grouping of hydrologically similar catchments? We utilize a heterogeneous dataset of 280 catchments located in the Eastern US to understand hydrologic similarity in a 6-dimensional signature space across a region with strong environmental gradients. Signatures are defined as hydrologic response characteristics that provide insight into the hydrologic function of catchments. A Bayesian clustering scheme is used to separate the catchments into 9 homogeneous classes, which enable us to interpret hydrologic similarity with respect to similarity in climatic and landscape attributes across this region. We finally derive several hypotheses regarding controls on individual signatures from the analysis performed here.

1 Introduction

Catchments provide a sensible (though not the only possible) unit for a hydrological classification system. Despite the degree of uniqueness and complexity that each catchment exhibits (Beven, 2000), we generally assume that some level of organization and therefore a degree of predictability of the functional behavior of a catchment exists (Dooge, 1986). This organization may be a result of natural self-organization or the co-evolution of climate, soils, vegetation and topography (Sivapalan, 2005). The uniqueness of catchments limits the success of hydrological regionalization, but the long-term use of statistical methods in hydrology suggests that some information transfer is possible. Hydrology has thus far not established a common catchment classification system that would provide order and structure to the global assemblage of these heterogeneous spatial units (McDonnell and Woods, 2004; Wagener et al., 2007) and which would provide a first order grouping of hydrologically similar catchments with implications for hydrological theory, observations and modeling (Gupta et al., 2008; McMillan et al., 2010).

Identifying and categorizing dominant catchment functions as revealed through a suite of hydrologic response characteristics, such as those extracted from observed streamflow-precipitation-temperature datasets, is one strategy to quantify the degree of similarity that may exist between catchments (McIntyre et al., 2005; Wagener et al., 2007; Oudin et al., 2008, 2010; Samaniego et al., 2010; Lyon and Troch, 2010; Haltas and Kavas, 2011). Understanding

how and why certain functional behavior occurs in a given catchment would ultimately shed new light on the reasons for the degree of similarity or dissimilarity that is exhibited between catchments (Gottschalk, 1985; Dooge, 1986). A range of benefits would be obtained if both catchment functions and their causes could be understood and formalized in a similarity framework, and therefore in a classification scheme (Grigg, 1965, 1967):

1. To give names to things, i.e., the main classification step.

2. To permit transfer of information, i.e., regionalization of information.

3. To permit development of generalizations, i.e. to develop new theory.

In the light of increasing concerns about non-stationarity of the responses of hydrologic systems (Milly et al., 2008; Wagener et al., 2010), we add a fourth benefit:

4 To provide a first order environmental change impact assessment, i.e., the hydrologic implications of climate, land use and land cover change.

All four of the above listed benefits are objectives of a catchment classification system to achieve order and new understanding while also providing predictive power. Achieving a generalization of knowledge beyond individual catchments or beyond a particular dataset has been a particular struggle in hydrology, as well as in other sciences related to the natural world (Beven, 2000; Harte, 2002). We believe that the task of catchment classification will be an essential element in this much hoped for generalization. So how should one define hydrologic similarity or dissimilarity in a catchment classification system? Past strategies for classification have largely focused on physical similarity (e.g., similarity in physical characteristics, or how the catchments look) or on similarity of some (narrowly defined) characteristic of the streamflow record (mainly based on flow regimes). Below we argue that both approaches fall short in achieving all of the four benefits listed above, and that the general idea of catchment function (Black, 1997; Sivapalan, 2005; Wagener et al., 2007) can bridge the gap between these strategies and help fulfill the needs of a more general classification system.

Previous studies have demonstrated the usefulness of empirical analysis of large datasets through clustering to identify homogeneous groups of hydrologically-relevant entities. Examples include Ramachandra Rao and Srinivas (2006) for catchments, Bormann et al. (1999) for hydrologic response units, Bormann (2010) for soil types, McNeil et al. (2005) for water bodies and Panda et al. (2006) for chemical water types. At the catchment scale, Winter (2001) introduced the idea of hydrologic landscapes, which are defined

on the basis of similarity of climate, topography and geology, assuming that catchments that are similar with respect to these three criteria will behave similarly in a hydrological sense. This approach clusters the USA into 20 non-contiguous regions using over 40 000 units of about 200 km^2 size (Wolock et al., 2004). In a similar manner, Buttle (2006) suggests that, within a particular hydro-climatic region, three factors should provide first-order controls on the streamflow response of catchments: (1) typology – hydrologic partitioning between vertical and lateral pathways, (2) topology – drainage network connectivity, and (3) topography – hydraulic gradients as defined by basin topography. Borman et al. (1999) applied a classification to hydrological quantities and physically based model (soil-vegetation-atmosphere-transfer-scheme) to examine the hydrologic behavior of catchments with respect to different ecotypes (hydropedotopes) of one catchment in Germany. Borman (2010) utilized a hydrologic classification system based on soil texture groupings, assuming soil to be a major control of hydrologic similarity. Similarly, Ramachandra Roa and Srinivas (2005) investigates catchments contained within Indiana using physical features (area, channel length, channel slope, etc.) in order to classify catchments as physically similar. These studies make the implicit assumption that the physical (climate and landscape) properties considered are the dominant controls on the "hydrologic behavior" of a catchment and are therefore sufficient to group catchments that are hydrologically similar. However, Merz and Blöschl (2009) showed that land use, soil types, and geology did not seem to fully define the process controls on catchment behavior when analyzing over 400 catchments in Austria. In addition, the uniqueness problem discussed above can lead to unexpected behavior of some catchments that is difficult to predict a priori without very detailed information about the catchment. Therefore, to fully permit (hydrologic) information transfer and to achieve a generalization of the relationships between catchment attributes, climate and hydrologic responses, an explicit quantitative assessment of such relationships (a mapping) is required and has to be tested, rather than an implicit one as used by Winter (2001) and Buttle (2006).

Alternatively, assessing similarity in terms of certain streamflow characteristics has been particularly popular in aquatic ecology, due to the importance of flow characteristics for aquatic habitats (e.g., Poff et al., 2006; Olden and Poff, 2003; Monk et al., 2007), and in regime studies (e.g., Haines et al., 1988; Bower et al., 2004; Moliere et al., 2009). However, these studies are typically not aimed at understanding the behavior of the catchment, including the causes of a particular regime occurring beyond climatic differences between regions.

Black (1997) introduced the idea of hydrologic function, defined as the actions of the catchment exerted on the precipitation it collects. Wagener et al. (2007, 2008) expanded this idea by viewing catchments as non-linear space-time filters, which perform a set of common hydrologic functions,

broadly consisting of partitioning, storage, and release of water. Partitioning is defined as the process whereby incoming precipitation is partitioned at the land surface into several components (e.g., infiltration, interception and surface runoff). Storage refers to the mechanisms by which incoming precipitation is held in temporary storage before its eventual release from the catchment (e.g., soil moisture, groundwater or interception). Release of stored water is defined as the pathway (and state) through which water ultimately leaves the catchment (e.g., evaporation, transpiration or surface runoff). Wagener et al. (2007) suggested that (to a degree) these functional characteristics should be revealed and hence observable in selected signatures of the catchment responses to precipitation input, i.e., in characteristics of the streamflow hydrograph, soil moisture and vegetation patterns, and other hydrologic variables. Different observed characteristics will enable a more or less detailed view of catchment function. Other observations, e.g., isotopic tracers, likely provide additional information on internal pathways than what could possibly be derived from streamflow data alone (e.g., McGlynn et al., 2003; Weiler et al., 2003; McGuire et al., 2005; Tetzlaff et al., 2009; Broxton et al., 2009). However, the limited availability of such tracer data makes it necessary to understand how far more generally available data such as streamflow can provide first-order insight.

This paper provides the first ever cluster analysis to be performed with respect to hydrologic similarity, derived from the notion of catchment function, across a large geographic region with strong environmental gradients. The objective is to understand how catchments group and whether hypotheses regarding controls on similarity can be generated. We use an empirical approach to cluster catchments based on hydrologic similarity as defined by six key signatures. None of the signatures themselves are novel, but their combined use to quantify hydrologic function, and hence hydrologic similarity, is. The choice of streamflow as output variable, with all its limitations as discussed above, means that while we can utilize many catchments, the similarity analysis is limited to a first-order classification. Some functional equifinality, i.e., a limited ability to uniquely characterize hydrologic function, will necessarily remain. The results are valid within the hydro-climatic and landscape characteristic gradients in our dataset. We use the clustering result to speculate on more general signature controls that will have to be generalized through additional study, e.g., using numerical models as used by Carrillo et al. (2011).

2 Study catchments and data

A total of 280 catchments, spanning the eastern half of the United States, were used in this study. Catchments range in size from $67 \, \text{km}^2$ to $10\,096 \, \text{km}^2$ (though only a few very large catchments are included), and show aridity indices (long-

term ratio of annual potential evapotranspiration to annual precipitation rates) between 0.41 and 3.3, hence representing a heterogeneous dataset (See further details in Supplement). The size of the catchments ensures that hillslope scale controls do not affect any of the signatures, which has been shown to decline beyond some $10 \, \text{km}^2$ catchment size (Robinson et al., 1995). Ecoregions are delineated based on climatic and land cover information. The ones found in our study region are type 1 eco-regions 5, 8 and 9, which are defined as Northern Forests, Eastern Temperate Forests, and Great Plains, respectively (Omernik, 1987).

Time-series data of daily streamflow, precipitation, and temperature for all catchments were provided by the MOPEX project (Duan et al., 2006). The catchments within this dataset are minimally impacted by human influences. Streamflow information within this dataset was originally provided by the United States Geological Survey (USGS), while precipitation and temperature were supplied by the National Climate Data Center (NCDC). A total of 10 hydrologic years of data was used (1949 to 1958) in order to calculate the signatures. This time period was assumed to be long enough to capture climatic variability, but short enough to not be affected by climatic trends. An analysis of the impact of trends on the classification is outside the scope of this paper. To ensure precipitation quality, the MOPEX dataset assumes a minimum acceptable precipitation gauge density within each catchment defined following the equation,

$$N = 0.6A^{0.3} \tag{1}$$

where N is the number of precipitation gauges and A is the area of the catchment (km^2) (Schaake et al., 2000). The use of this guideline provides mean areal precipitation estimates at each time step and should result in less than 20 % error 80 % of the time (Schaake et al., 2006). The MOPEX dataset has been used widely for hydrologic model comparison studies (see references in Duan et al., 2006).

3 Signatures

Signatures quantify characteristics of the hydrologic response that provide insight into the functional behavior of the catchment. In this paper we will limit ourselves to signatures derived from widely available time-series information such as streamflow, precipitation, and temperature as the basis of a first-order analysis. These key signatures were chosen, from a much larger list of possible indices (see list in Yadav, 2007), by selecting those that are largely uncorrelated and that have an interpretable link to catchment function as discussed in the next sections. None of the chosen signatures scales with catchment size. The chosen signatures are: runoff ratio, baseflow index, snow day ratio, slope of the flow duration curve, streamflow elasticity, and rising limb density. In the remainder of this section we will provide a brief definition of each of the six signatures.

3.1 Runoff ratio

Runoff Ratio (R_{QP} [-]) is defined as the ratio of long-term average streamflow, Q, to long-term average precipitation, P,

$$R_{QP} = \frac{Q}{P} \tag{2}$$

It represents the long-term water balance separation between water being released from the catchment as streamflow and as evapotranspiration (assuming no net change in storage) (Milly, 1994; Olden and Poff, 2003; Poff et al., 2003; Sankarasubramanian et al., 2001; Yadav, 2007). A high runoff ratio identifies a catchment from which a large amount of water exits as streamflow (streamflow dominated or blue-water dominated), whereas a low value of runoff ratio identifies a large amount of water exiting the catchment as evapotranspiration (ET dominated or green-water dominated).

3.2 Slope of the Flow Duration Curve

The Flow Duration Curve (FDC) is the distribution of probabilities of streamflow being greater than or equal to a specified magnitude. FDC is typically derived from hourly or daily (and sometimes monthly) streamflow data (e.g., Vogel and Fennessy, 1994; Jothityangkoon et al., 2001; Jencso et al., 2009). To quantify an index of flow variability, the slope of the FDC (S_{FDC} [-]) is calculated between the 33rd and 66th streamflow percentiles, since at semi-log scale this represents a relatively linear part of the FDC (Yadav et al., 2007; Zhang et al., 2008). A high slope value indicates a variable flow regime, while a low slope value means a more damped response. Damped response can arise as a result of a combination of persistent (wide-spread and year-round) rainfall and/or the dominance of groundwater contribution to streamflow. The signature is defined as,

$$S_{FDC} = \frac{\ln(Q_{33\%}) - \ln(Q_{66\%})}{(0.66 - 0.33)} \tag{3}$$

where S_{FDC} is the slope of the flow duration curve, $Q_{33\%}$ is the streamflow value at the 33rd percentile, $Q_{66\%}$ is the streamflow value at the 66th percentile.

3.3 Baseflow index

Base Flow Index (I_{BF} [-]) is the ratio of long-term baseflow to total streamflow (e.g., Arnold et al., 1995; Vogel and Kroll, 1992; Kroll et al., 2004). A high value of I_{BF} defines a catchment with higher baseflow contribution, i.e., more water using long flowpaths through the catchment. A range of algorithms has been proposed and compared to perform a separation of quick flow and baseflow from observations of streamflow alone (Kroll et al., 2004; Eckhardt, 2005; Institute of Hydrology, 1980; Arnold et al., 1995; Arnold and Allen, 1999; Wittenberg and Sivapalan, 1999; Laaha and Blöschl, 2006). In this study we use the one-parameter single-pass digital filter method (DFM) based on previous studies reported by Arnold et al. (1995) and Lim et al. (2005). We do not consider the specific choice of filter crucial in this study since we focus on the relative differences between catchments, rather than on the actual values. The filter applied is defined as follows,

$$Q_{Dt} = c Q_{Dt-1} + \frac{1+c}{2}(Q_t - Q_{t-1}) \tag{4}$$

where $Q_{D,t}$ is the direct flow value at time-step t, Q_t is the total flow at time step t, and c is a parameter. The parameter c was set at a value of 0.925 based on a comprehensive case study performed by Eckhardt (2007). The value of the baseflow $Q_{B,t}$ at time-step t is then given by,

$$Q_{Bt} = Q_t - Q_{Dt} \tag{5}$$

The baseflow index is therefore,

$$I_{BF} = \sum \frac{Q_B}{Q} \tag{6}$$

where the summation is carried out over all time steps of the study period.

3.4 Streamflow elasticity

Streamflow elasticity (E_{QP}[-]) describes the sensitivity of a catchment's streamflow response to changes in precipitation at the annual time scale. It can be calculated by taking the inter-annual difference between annual streamflow divided by the inter-annual difference between annual precipitation, which is then normalized by the long-term runoff ratio (Schaake, 1990; Dooge, 1992). Based on the study by Sankarasubramanian et al. (2001) we assume that the median value provides a more stable value of this index, i.e.,

$$E_{QP} = \text{median}\left(\frac{dQ}{dP}\frac{P}{Q}\right) \tag{7}$$

where E_{QP} is the streamflow elasticity, dQ (dP) is the difference between the previous year's streamflow (precipitation) and the current year's streamflow (precipitation), P is the mean annual precipitation, and Q is the mean annual streamflow. The median value of E_{QP} is considered as a robust measure, since it filters out outliers, which may significantly affect the mean value (Sankarasubramanian et al., 2001; Sankarasubramanian and Vogel, 2003). E_{QP} is the percentage change in streamflow divided by the percentage change in precipitation. A value of 1 indicates that a 1 % precipitation change leads to a 1 % change in streamflow. A value greater or less than 1 would, respectively, define the catchment as being elastic, i.e., sensitive to change of precipitation, or inelastic, i.e., insensitive to a change of precipitation.

3.5 Snow day ratio

The snow day ratio (R_{SD} [-]) is defined as the number of days that experience precipitation when the average daily air temperature is below 2 °C, divided by the total number of days per year with precipitation. This value provides an indicator of the amount of precipitation that falls and is stored as snow. It can be related to the seasonality of the catchment response (Woods, 2003). The ratio is defined as,

$$R_{SD} = \frac{N_S}{N_P} \tag{8}$$

where N_S is the number of days with precipitation and a daily average temperature below 2 °C, and N_P is the number of days with precipitation. A high value of R_{SD} suggests more snow storage with a significant impact on the intra-annual variability of streamflow.

3.6 Rising limb density

The sixth signature considered in this study is called Rising Limb Density (R_{LD}). R_{LD} describes the flashiness of the catchment response and is defined as the ratio of the number of rising limbs (N_{RL}) and the total amount of time the hydrograph is rising (T_R) (Morin et al., 2002; Shamir et al., 2005). The equation is given as,

$$R_{LD} = \frac{N_{RL}}{T_R} \tag{9}$$

R_{LD} is a descriptor of the hydrograph shape and smoothness without consideration for the flow magnitude. A small the signature value indicates a smooth hydrograph.

4 Methods: cluster analysis

Cluster analysis is the process of grouping similar entities (catchments) according to one or more chosen similarity measures (signatures in our case), while concurrently separating those entities that are different. There are three common types of clustering algorithms: agglomerative hierarchical clustering, k-means clustering, and fuzzy partition clustering. All three strategies of unsupervised clustering require some subjective choices that define the clustering process, e.g., the distance metric used, and there is consequently not one single solution to this kind of analysis. The objective in this study is therefore to use an empirical analysis to investigate how the similarity between catchments defined by the six signatures might create groupings, and is not to derive at an ultimate classification result, which would always depend on the choices we made and the dataset used anyway. To account for the uncertainty in the classification process, we used a fuzzy partitioning algorithm that enables us to analyze the uncertainty in the resulting classification.

The method chosen for this study is a fuzzy partitioning Bayesian mixture clustering algorithm implemented in the AutoClass C software package (version 3.3.4) (Stutz and Cheeseman, 1995; Cheeseman and Stutz, 1996; Archcar et al., 2009; Kennard et al., 2010). Bayesian mixture modeling is a probabilistic approach in which marginal likelihoods for different classification realizations are estimated and ranked against all other realizations. The classification with the highest posterior probability is ultimately chosen as the most likely realization (Webb et al., 2007). Each catchment is therefore assigned to a particular class with a certain probability, called here the probability of class assignment. The number of classes is automatically decided during the classification process. A catchment could be allocated to different classes due to the probabilistic nature of the algorithm, and it is only the primary class assignment that is listed, which pertains to the class assignment with the highest probability. The input variables characterizing the catchments, i.e., the signatures, were log transformed and modeled as normally distributed continuous variables with an associated degree of uncertainty. Additionally, these variables are scaled such that the magnitude differences between signatures do not cause any additional weighting in the calculation of distance metric. The output of the clustering process is analyzed with respect to the probability of each catchment being member of a particular class, the class strength (calculated as a heuristic measurement, where a high class strength means a narrow range of signature values), and the importance of each signature in separating the different classes (calculated using the Kullback-Leibler distance metric (Kullback and Leibler, 1951)). Another advantage of the clustering algorithm used here is the ability to consider correlation between signatures. We account for the covariate information from two correlated similarity measures, in our case signatures SFDC and IBF, using a multi-normal model. We therefore chose this particular algorithm due to its ability to consider uncertainty and correlation in contrast to other clustering strategies (Jain et al., 1999).

Due to the probabilistic nature of the AutoClass-C algorithm, classification realizations will change over multiple runs. To test for the stability of the results across these different realizations, we use the Adjusted Rand Index (ARI, Rand, 1971; Hubert and Arabie, 1985), which takes a value of 0, if the agreement between two classification schemes is no better than mere chance, and 1, indicating perfect agreement between the two classification schemes. We use ARI to test the similarity of classification results when the algorithm is initiated multiple times.

5 Results and discussion

5.1 Signature relationships and spatial variability

Figure 1 shows the relationships between the signatures both visually and numerically. In addition to the linear coefficient of correlation, C_{Lin}, the Spearman rank correlation
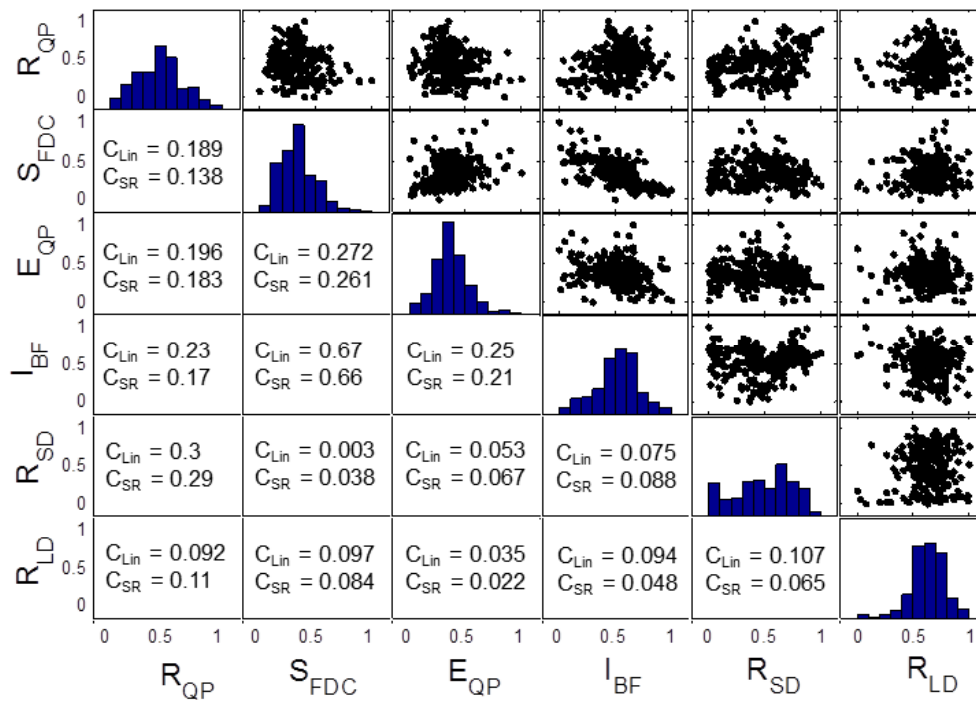
Fig. 1. Distributions of individual signatures shown as histograms and correlations between the signatures shown as scatter plots as well as numerical values. Correlation (C) is calculated using linear correlation (Lin) and Spearman Rank (SR) correlation coefficients, both ranging from zero to one.

coefficient, C_{SR}, has also been calculated to show potential non-linear relationships. In the set of signatures used, Baseflow Index (I_{BF}) and Slope of the Flow Duration Curve (S_{FDC}) show the highest linear correlation (0.67). While the correlation between these two signatures is partially created by a smaller number of very high S_{FDC} values, the correlation is considered during clustering as discussed in the methods section.

The different spatial patterns that the six signatures produce across the study domain can be seen in Fig. 2. Runoff ratio, RR, shows high values in the humid region along the Appalachian mountain and connected plateau regions, which decrease with increasing distance from this area, especially towards the central US (see also Sankarasubramanian and Vogel, 2003). Figure 2b shows that the smallest values of the slope of the FDC are located on the southeastern side of the Appalachian mountain range and west of this area. Values of streamflow elasticity and rising limb density show much greater heterogeneity than the first two signatures (Fig. 2c and f). High values of baseflow index can be found along the Eastern coastal US and around the Great Lakes region (Fig. 2d), where more permeable soils and bedrock dominate (Wolock et al., 2004; Santhi et al., 2008). Values decrease when moving towards the east where soils and bedrock are more impermeable. As expected, the ratio of snow days correlates highly with latitude (Fig. 2e). These spatial patterns further underline the relative independence of the different

signatures and attest to their suitability for the similarity analysis. Figure 2 also shows that, taken individually, there are strong regional patterns in the variations of several of the signatures (Runoff Ratio, Ratio of Snow Days, Baseflow Index, Slope of FDC). The variability in the spatial patterns seen also suggests a difference in the variables that control the signatures. Some characteristics change slowly in space (e.g., Runoff Ratio), which is likely due to the spatially slowly changing climate. Other signatures show much more abrupt changes, which suggests controls related to soil or geological characteristics.

5.2 Cluster analysis

The cluster analysis was applied to all 280 catchments within the six-dimensional signature space. The analysis aimed at addressing the following questions: (1) how do the catchments group with respect to the signatures used? (2) What spatial patterns of clusters emerge? (3) What hypotheses regarding the function of catchments and what physical or climatic controls on this functional behavior can be derived?

The cluster analysis identified 9 different classes of varying size for the most likely classification. The classification process was repeated 20 times and the Adjusted Rand Index (ARI) between classification schemes of 15 of the 20 runs was above 0.90. Furthermore, 7 of these 20 runs were found to be nearly identical and one of these 7 runs was therefore
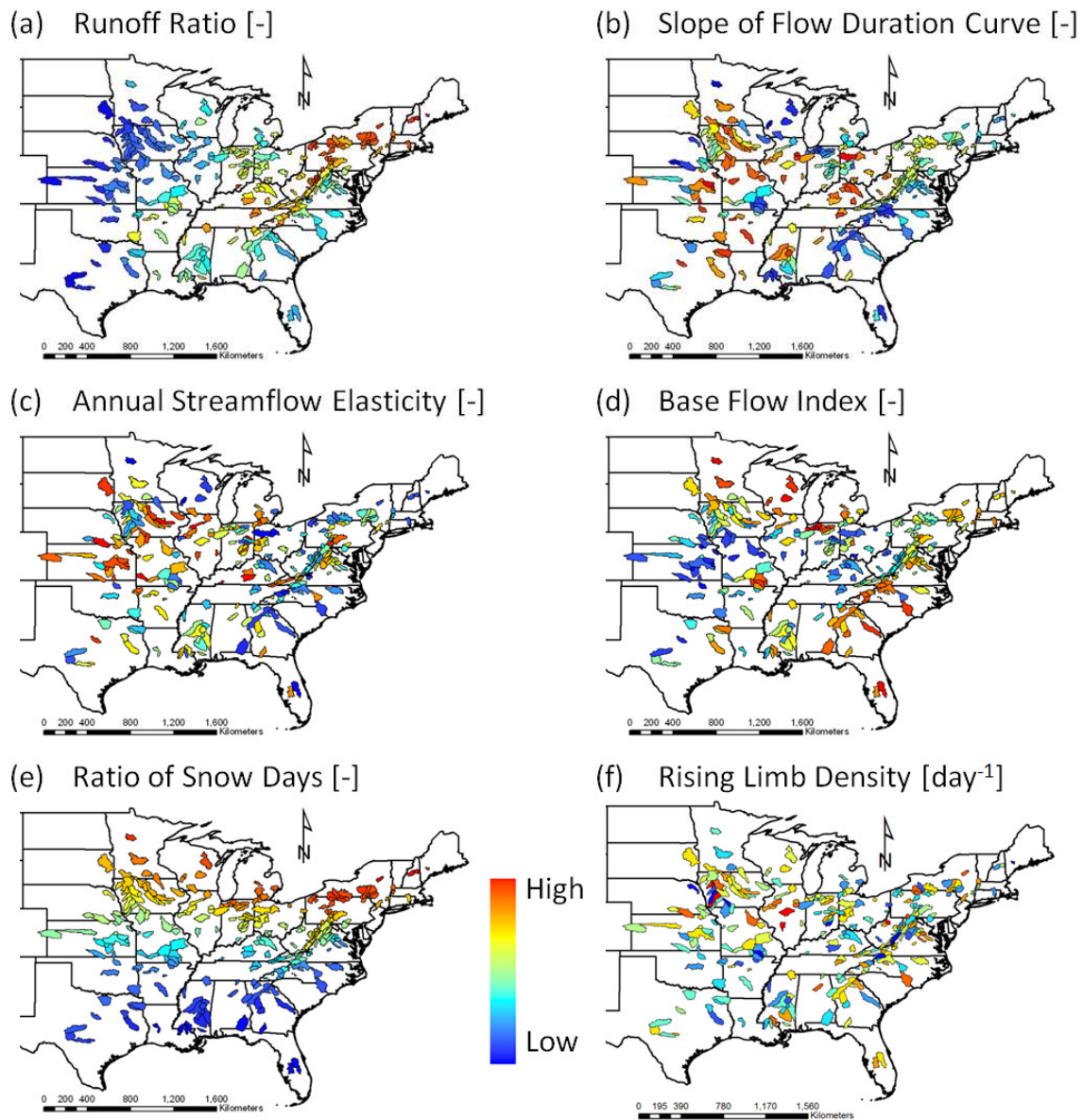
Fig. 2. Maps showing the spatial distribution of signatures for each of the study catchments. The color of the catchment corresponds to the high (red) and low (blue) values, as shown by the colorbar. Plots show spatial distributions of: **(a)** mean annual runoff ratio (R_{QP}). **(b)** Slope of the FDC (S_{FDC}). **(c)** Streamflow elasticity (E_{QP}). **(d)** Base flow index (I_{BF}). **(e)** Ratio (or fraction) of snow days (R_{SD}). **(f)** Rising Limb Density (R_{LD}). The actual range in values can be found in the Supplement.

used as the final classification. Results were also screened for extremely small or large classes, and for generally providing high probability of class memberships to ensure that no unreasonable solutions were used.

The heuristic measures describing the algorithm and classification performance (discussed in the methods section) are visualized in Fig. 3 for the chosen result. Figure 3a shows a histogram of the probabilities that a catchment belongs to the assigned class. The histogram indicates that the vast majority of catchments is assigned with probabilities above 0.9 and hence that they are very likely classified correctly. The num-

ber of catchments per class varied between 5 and 82 (Fig. 3b). All classes show a relatively high class strength (Kennard et al., 2010), i.e., the variability of signature values within each class is rather low. Classes 2 and 3 have the highest values, while those for class 5, 8, and 9 are somewhat lower.
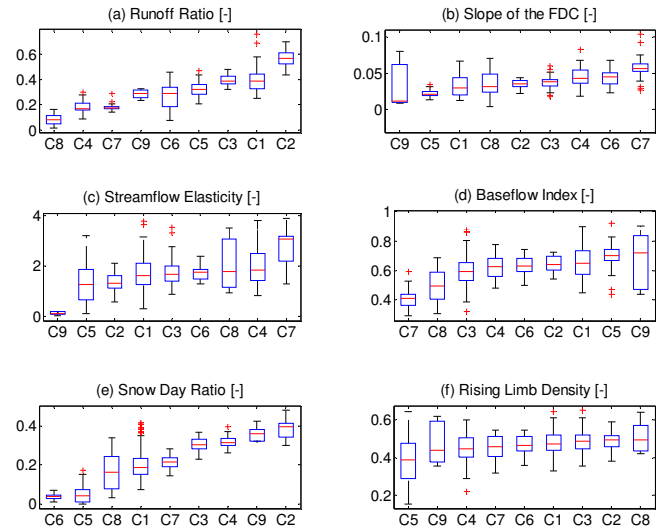
The relative value of attribute influence of each signature describes the contribution of each signature to the classification. This measure represents the separation of classes due to each signature, and is calculated from the average Kullback-Leibler distance between attribute distributions in individual classes and the overall distribution found in the
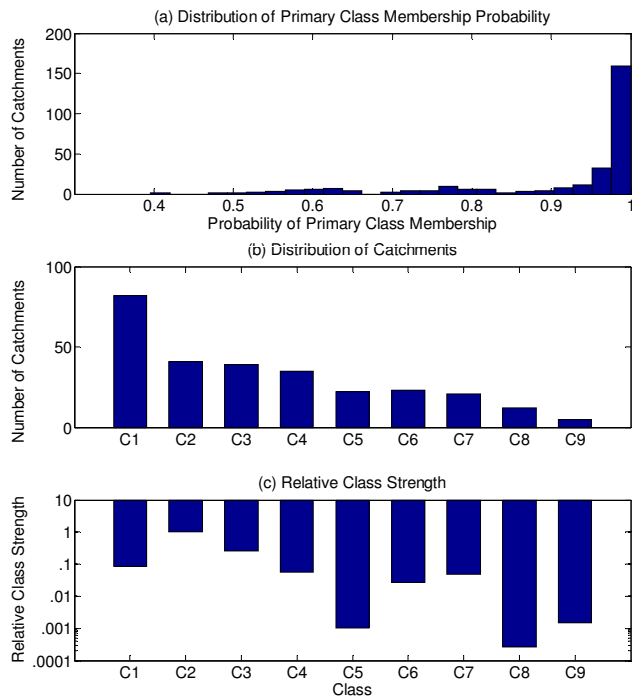
Fig. 3. (a) Distribution of the primary class membership probability for all 280 catchments. (b) Histogram of the number of catchments within each class. (c) Relative class strength. Measure of how similar catchments are in a particular class normalized to the strongest class (value of 1 is the strongest value, all other classes are less).


Fig. 4. Distribution of signature values for each class. Classes are sorted by median value from low to high (left to right). (a) Mean annual runoff ratio (R_{QP}). (b) Slope of the FDC (S_{FDC}). (c) Streamflow elasticity (E_{QP}). (d) Base flow index (I_{BF}). (e) Ratio (or fraction) of snow days (R_{SD}). (f) Rising Limb Density (R_{LD}).

full dataset (Webb et al., 2007). The attribute influence increases as the variance between the signature means of each class increases. Its values range from 1 (highest contribution) to 0 (no contribution). The order of signature influence on the clustering result was: Streamflow Elasticity (1), Ratio of Snow Days (0.98), Runoff Ratio (0.862), Slope of the FDC (0.462), Baseflow Index (0.462), and Rising Limb Density (0.201). These values suggest that all the signatures provided information for the classification, though RLD was not very influential. It also suggests that mainly climate-controlled signatures dominate the classification, further adding to the evidence supporting the dominant role climate has in controlling catchment behavior (see also Rosero et al., 2010).

Cluster results regarding the distribution of signature values in each cluster are shown as a box and whisker plots in Fig. 4. The catchments in Fig. 4 are sorted from left to right by increasing median value of the signature shown. The spatial distribution of the clusters is shown as a map with a corresponding heatmap in Fig. 5. The heatmap shows the distribution of signature values indicated by colors for all catchments. Figures 6 and 7 are used to analyze possible controls on the signature-based classes identified. Geographic location is used to structure the discussion of the resulting classification, starting in the Northeastern US.

Catchments in the northeastern United States (class C2) are characterized by high ratios of streamflow to precipitation (high RQP) and large amounts of snow (high RSD) (Fig. 4). These catchments are located in humid continental climate with low energy availability and hence low evapotranspiration. Snow storage is important in controlling seasonal variability of runoff, i.e., these catchments have the highest ratio of snow days in the dataset. Catchments located in class C2 can be found in an area ranging from Maine to Pennsylvania. These catchments have the smallest basin sizes in the dataset with the highest slope (max SLOPE and min DA), with long and frequent storms (max SD and NP), and the lowest maximum temperatures (TMAX). This class consists mainly of catchments with a very low precipitation seasonality index (PSI; Figs. 6 and 8), meaning that precipitation amounts are distributed relatively evenly throughout the year (a uniform distribution would be 0). PSI values are generally low for the eastern US (less than 0.6) and lowest in the Northeastern US, compared to the southwestern US where rain falls mainly in winter (Pryor and Schoof, 2008). All catchments in this class are energy limited with the smallest aridity indices in the study region (AI = PE/P below 1), suggesting that these catchments are controlled by low average energy availability throughout the year, but with strong seasonal variability, as well as a relative uniform distribution of moisture input in time.

Catchments slightly further to the south (in Pennsylvania and Virginia), cluster C3, extend westward to Indiana with lower runoff ratios (median of 40% versus about 55% for class C2) and less dominant snow storage, while the values
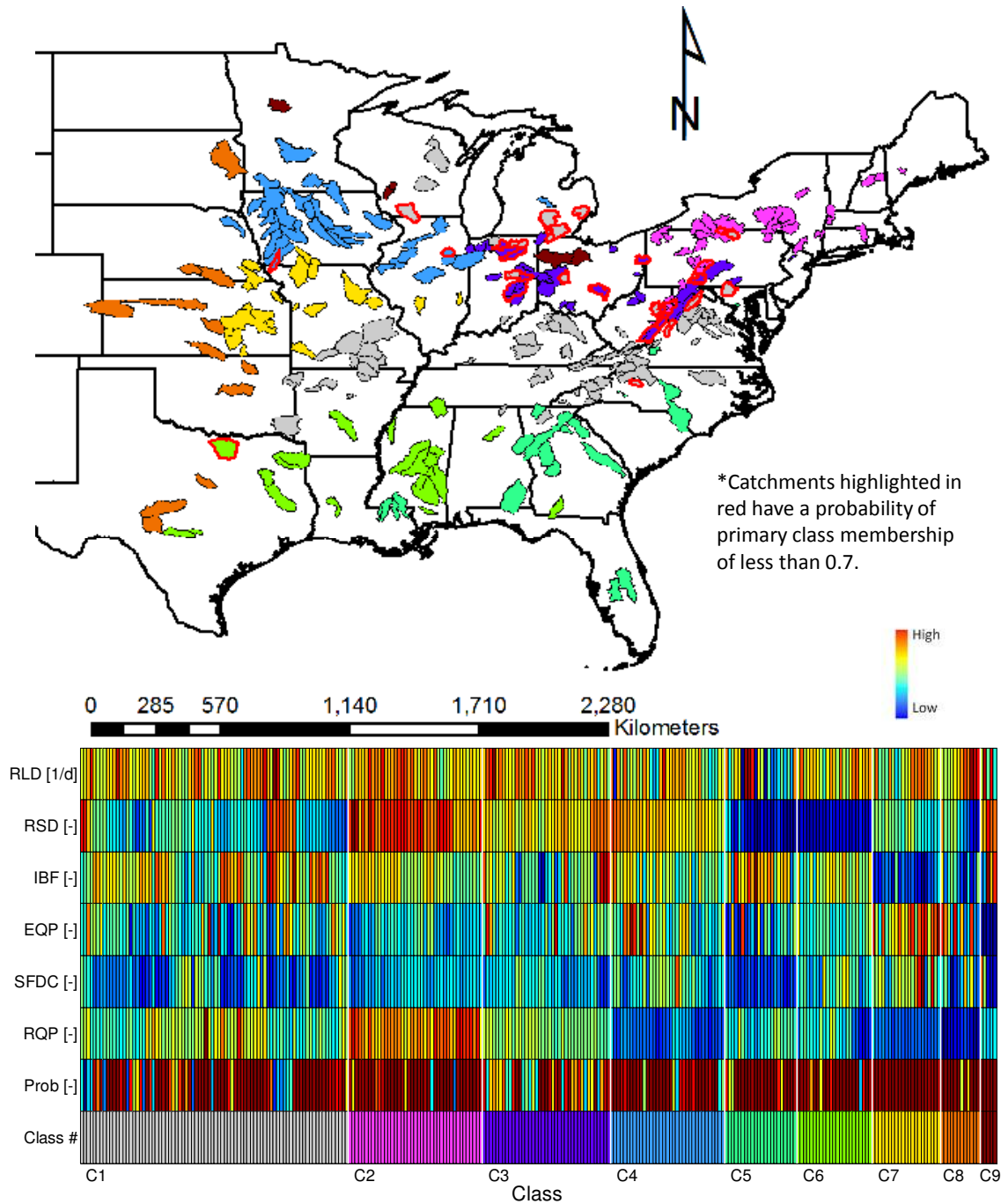
Fig. 5. The top figure represents the spatial distribution of catchments according to their classes. The catchments are color coded according to Class # as shown in the lower part of the recursive pattern plot below (this plot type is also sometimes called a heat map). The first 6 rows of the recursive pattern plot show signature values (high values shown as red, low being blue). The 7th row indicates the probability of a catchment's primary class assignment. The 8th and last row represents the class color code used the in the map above.

of SFDC stay relatively similar to those catchments further north (Fig. 4). There are generally strong similarities in physical and climatic characteristics between the catchments of classes C3 and C2: they are the smaller catchments, with a high fraction of poorly drained soils (high HGC), the longest storm durations and the lowest aridity indices. Differences to the previous class are mainly topography and land use, since C3 catchments are at a lower elevation, with less slope, and have more agricultural land use and therefore lower root zone depths (Figs. 6 and 8).
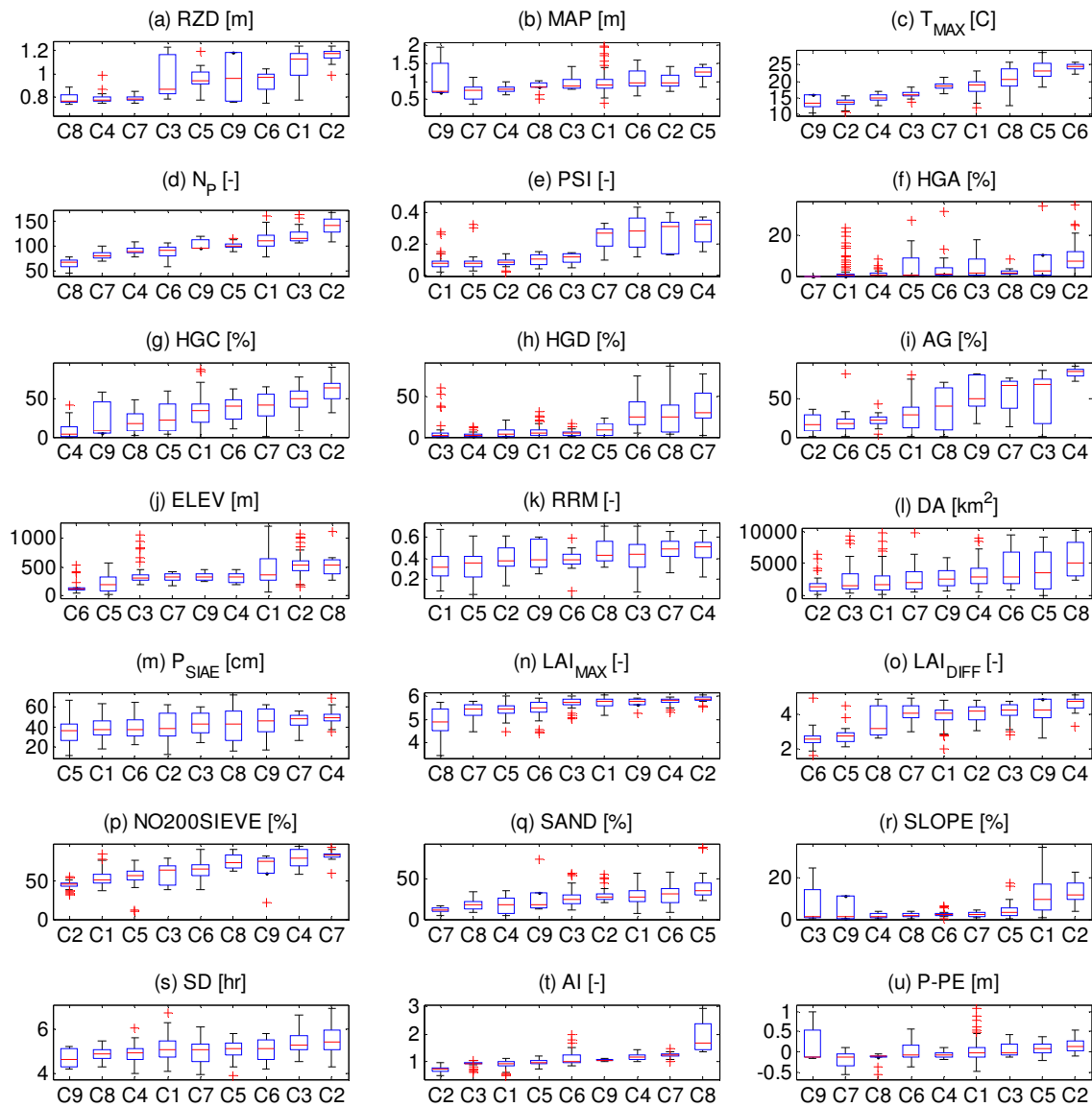
Fig. 6. Box-Whisker plots of the clusters with respect to landscape and climatic characteristics. Explanations of abbreviations can be found in Table 1.

Within the extent of C3 catchments, a small collection of catchments from C9 can be found in Northwest Ohio. This class is separated by very low streamflow elasticity values, along with the smallest variability of this signature in any class. Class 9 does show the lowest median value of Slope of the FDC, and the highest median value of Baseflow Index, however the variability of these signatures within this class is the largest of any class. Class 9 also shows a larger amount of snow than C3, suggesting a climatic separation between these regions. Climatically, C9 experiences the lowest temperatures and shortest storm duration in the dataset, along with some of the highest precipitation seasonality found within our dataset.

Continuing southeastward down the Appalachian Mountain entering North Carolina, we witness a decrease in the values of Ratio of Snow Days in class C1, although the variability of this signature increases in this class, which also contains the largest number of catchments. We also see a small decrease in the Slope of the FDC (and conversely, a slightly higher baseflow index). Catchments belonging to this class are spread along the same latitude and are mainly separated from classes C2 and C3 by a lower snow day ratio attributable to their location (median of about 20 %, Fig. 4). From Fig. 6, we can also see that these catchments experience low precipitation seasonality indices, a low fraction of poorly drained soil along with high percentage of sand, and hence high soil permeability, as well as the lowest relief

**Table 1.** Explanation of physical and climatic properties used.

Variable	Description [units]	Minimum Value	Maximum Value
AG	Fraction of land used as agriculture within 800 m of the stream [-]	0.04	91.87
AI	Aridity Index, defined as the ratio between mean annual Potential Evapotranspiration (PE) and Precipitation (P). PE was calculated using a modified Hamon algorithm (Dingman, 2002) [-]	0.490	2.95
DA	Contributing drainage area [km^2]	67	10096
ELEV	Mean elevation of the catchment [m]	21.9	1212
HGA	Percentage of soil within the "A" Hydrologic Group (well drained soiled). Soils are deep and well drained and, typically, have high sand and gravel content [%]	0	35.04
HGC	Percentage of soil within the "C" Hydrologic Group (poorly drained soiled). The soil profiles include layers impeding downward movement of water and, typically, have moderately fine or fine texture [%]	0	91.77
HGD	Percentage of soil within the "D" Hydrologic Group (very poorly drained soiled). Soils are clayey, have a high water table, or have a shallow impervious layer [%]	0	88.01
LAI_{DIFF}	Difference between the maximum Leaf Area Index (LAI_{MAX}) and the minimum Leaf Area Index (LAI_{MIN}) based on vegetation type [-]	1.619	5.146
LAI_{MAX}	Maximum Leaf Area Index based on vegetation type [-]	3.46	6.04
MAP	Mean annual precipitation [mm yr^{-1}]	46.6	207.2
NO200SIEVE	Percent soil passing through the No. 200 sieve. [%]	10.5	94.4
NP	Number of days with measurable precipitation per year [days]	46.7	167
P-PE	The difference between mean annual Precipitation and Potential Evapotranspiration. Used in hydrologic landscapes and included as a reference [m]	-0.564	1.083
P_{SIAE}	Capillary fringe height [cm]	12.2	71.36
PSI	Precipitation seasonality index, as defined by $PSI = \frac{1}{MAP} \sum_{12}^{n=1} |X_n - \frac{MAP}{12}|$. A high value means that precipitation is seasonal, and a low value means that the precipitation is uniformly distributed throughout the year. X_n is the monthly precipitation for month n [-]	0.017	0.434
RRM	Relief ratio ($Elev_{median}$-$Elev_{min}$)/($Elev_{max} - Elev_{min}$) [-]	0.059	0.705
RZD	Mean rootzone depth [m]	0.732	1.243
TMAX	Mean maximum monthly temperature within the dataset [°C]	10.4	28.7
SAND	Percentage of sand found within the catchment. Used as a metric in hydrologic landscape regions and included as comparison [%]	4.60	89.45
SLOPE	Mean slope found within the catchment. Used as a metric in hydrologic landscape regions and included as comparison [%]	0.10	34
SD	Mean Storm duration, defined as when the precipitation starts and zero, increases, and decreases back to zero [hours]	3.88	6.93

ratio (RRM) of all classes. Classes C1 and C5 (the next class further south) are the baseflow dominated catchments in the dataset.

In the catchments further south of class C1, i.e., cluster C5, we find a persistence of high IBF and low SFDC values. The "number of snow days' " signature decreases even further, suggesting a climate-based separation between the classes C1 and C5. Catchments in C5 also have some of the lowest elasticity values (Fig. 4c), which is in line with their low flow duration curve slope (suggesting high storage in the catchment). The higher baseflow fraction creates a smaller streamflow response to changes in precipitation. Catchments

within this class experience high maximum temperatures and large amounts of rainfall (>1200 mm yr^{-1}).

West of class C5, in the southern Mississippi river basin, are catchments of class C6. While climatically similar, e.g., similarly water limited with negligible snow as class C5, the baseflow index decreases, with decreasing percentage sand levels and deeper mean root zone depths (Fig. 6). This class is also quite similar to the catchments grouped in class C1, even though located further north, with respect to geologically controlled signatures (e.g., I_{BF}), but with lower runoff ratio.
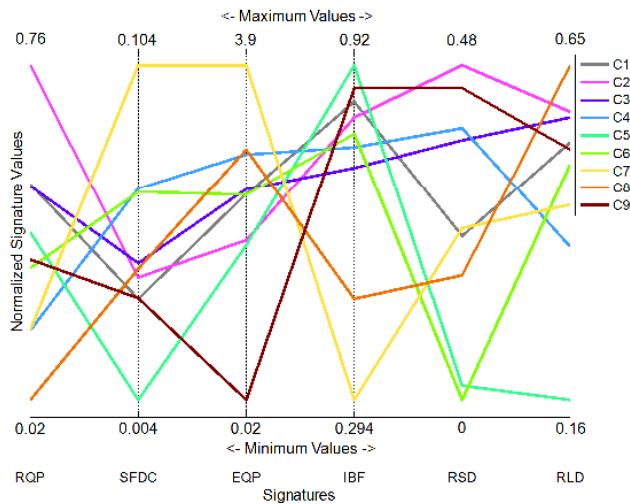
Fig. 7. Parallel coordinate plot in which the median signature values of the individual clusters are connected. This plot visualizes relationships between the signature values of the different clusters. E.g., inverse relation between EQP and IBF can be identified by the cluster lines crossing each other.

The catchments located furthest west in the study region (belonging to classes C8, C7 and C4) are characterized by the lowest runoff ratios. Generally less than 20 percent of precipitation is released from the catchments as streamflow (Fig. 4a). These are water-limited catchments (AI > 1) that receive less precipitation than the other areas of the study region, while the fraction of snow increases when moving from south to north (C8 to C7 to C4) (Fig. 4e). The catchments of this class, along the western boundary of the study region, are approaching a more arid area of the Köppen classification system (Peel et al., 2007). This area also exhibits a high precipitation seasonality index, i.e., a non-uniform distribution of precipitation through the year. These catchments are located in areas that are primarily cultivated land use, demonstrated by a very higher percentage of agriculture found within 800 m of the stream (riparian zone) (Figs. 6 and 8). Catchments south of Iowa, class C7, exhibit the highest flow duration curve slopes, SFDC, and elasticity values, EQP, as well as the lowest baseflow indices in the dataset. The latter is likely related to the very low percent of soils classified as sand, and the highest fraction of very poorly drained soils.
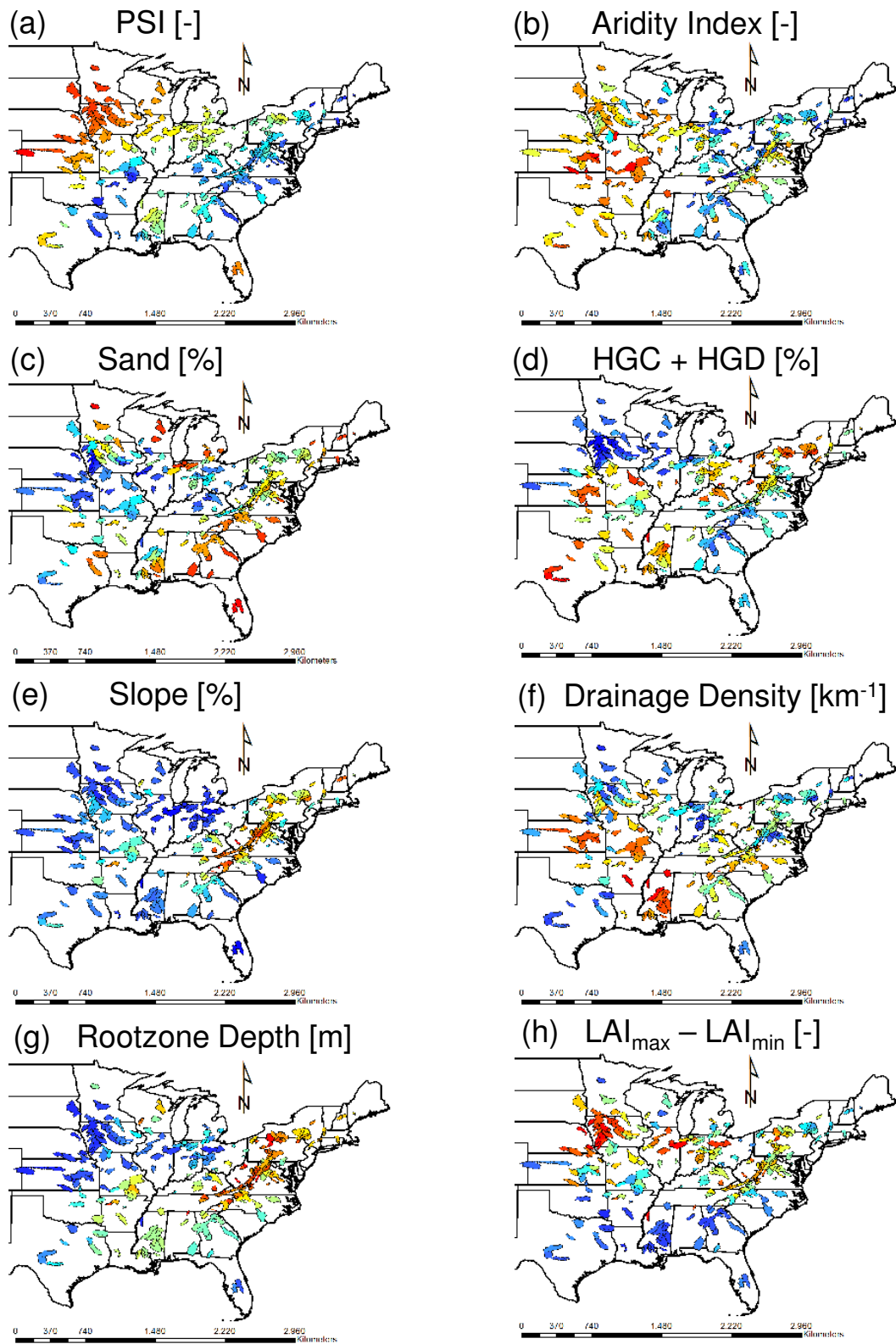
5.3 Discussion

Overall it seems that signatures, which vary along a climatic gradient (RQP, EQP, RSD) are exerting a stronger control on separating the catchments into different classes than the signatures that are likely to be more impacted by topographic, geological and land cover variability (IBF, SFDC, RLD). This highlights one problem with this type of empirical analysis in which the result is controlled by the particular

gradients present in the analyzed dataset, hence making it difficult to generalize beyond the data at hand. This result further suggests that a general regionalization of signatures across the region might not be the best strategy for some of the signatures, but that the region has to be broken up into smaller subregions (see the example by Laaha and Blöschl, 2006). The degree of equifinality of controls might also be reduced if further variables characterizing the functional behavior of the catchments would be included. For example, one would expect tracer data to provide a better separation of flowpaths/residence times and hence enable a refinement of the hydrologic function of the catchments.

This conclusion was not unexpected and, as stated on the outset of this paper, no cluster analysis can produce a general classification system because the results are depending on the dataset used and subjective decisions made (mainly choice of algorithm and distance metric). However, the clustering results help to understand controls across the study region, and potentially enable us to derive a small number of hypotheses. One should analyze these hypotheses further in a more idealized setting (i.e., not empirical) to understand the generality of the results found here. Multiple authors have advocated the use of "virtual experiments" for this purpose, i.e., by analyzing modeled or synthetic realities rather than actual systems (e.g., Bashford et al., 2002; Weiler and McDonnell, 2004; Winter et al., 2004; van Werkhoven et al., 2008). So what does the empirical analysis above suggest? The main issue we focus on is the suggested variability of controls on similar hydrologic signatures, and hence on hydrologic function in the context of this paper.

Streamflow elasticity with respect to precipitation is modified by the permeability characteristics of a catchment. Our results suggest that high elasticity values (clusters 8 and 7) relate to low I_{BF} values and vice versa (clusters 9, 5, 2, 1) (Figs. 4 and 7). Cluster 8 and 7 have low % Sand and the highest percentage poorly drained soils (HGD), and hence the smallest potential for buffering precipitation variability. Clusters 5, 2 and 9 have high % Sand, and 5 and 2 also have a high percentage well drained soils (HGA), and therefore a high potential for buffering. This result is similar to the conclusions of Sankarasubramanian and Vogel (2003) who analytically derived a parameter (parameter b of the abcd model) they refer to as soil moisture holding capacity, which they found to buffer streamflow variability and that they consider regionalizable using soil permeability. The classes with the highest elasticity values (8, 4, 7) are also the classes with the shallowest roots (lowest RZD) and the lowest runoff ratio (highest fraction of evaporated precipitation). Class C2 on the other hand has the highest root zone depth and the highest runoff ratio (highest of precipitation becoming streamflow). This interaction between climate, soils and vegetation is also shown in the five-dimensional plot of Fig. 9. It shows that deep-rooted vegetation coincides with high runoff ratios (energy limited catchments), but only if the catchments have mainly poorly or very poorly drained soils. Results like these

Fig. 8. Spatial map of selected climatic and physical characteristics of the study catchments.

www.hydrol-earth-syst-sci.net/15/2895/2011/

Hydrol. Earth Syst. Sci., 15, 2895–2911, 2011
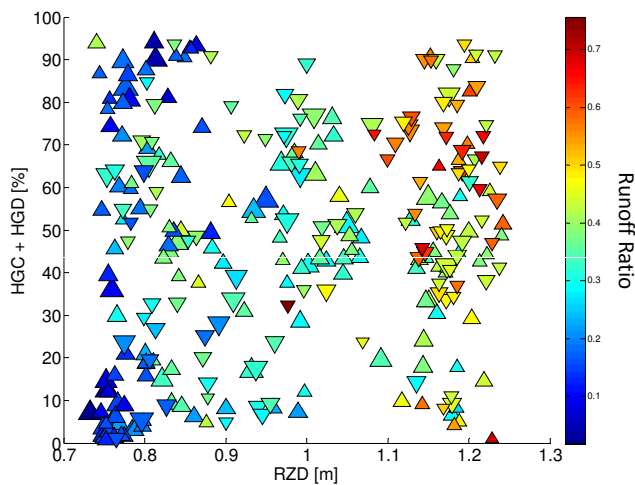
Fig. 9. Five-dimensional plot of signature runoff ratio versus soil, vegetation and climate characteristics. Triangles are color coded by Runoff Ratio. Size of triangles is proportional to drainage area (large point means large value). Triangles pointed upwards mean water limited catchments (pointed down means energy limited [AI > 1], upwards means water limited).

hint at a co-evolution of soil-climate-vegetation, which is further explored numerically in the parallel classification study by Carillo et al. (this issue).

Spatial proximity is a valuable first indicator of hydrologic similarity because it reflects the strong climatic control on catchment behavior, which varies slowly in space. Many researchers have commented on the value (Merz and Blöschl, 2004, 2005; Parajka et al., 2005; Oudin et al., 2008) or lack of value (for example with respect to drought characteristics, see Tallaksen and van Lanen, 2004) of catchment spatial proximity in predicting hydrological similarity. The empirical results shown here suggest that spatial proximity clearly plays an important role as an indicator of similarity. However, the results also suggest that spatial proximity is generally reflecting similarity in other characteristics. The different clusters show strong spatial connectedness (clusters 4, 7, and 2), show large patches with "outliers" (6, 5, 3, 8, 9), or are relatively widely distributed (1 and 6). Combining Figs. 4, 5 and 7, we can see that cluster 4 is a connected group of catchments with very low (high) values of runoff ratio (ratio of snow days), and with very little variability in both signatures. An analysis of Fig. 6 shows little variability and extreme (within the dataset) values in landscape (lowest root zone depth, highest % HGC and HGD) and climatic (lowest MAP, highest PSI) characteristics for this cluster. This result suggests that the catchments in this cluster are very different from the rest. Another connected cluster is C2 in the northeastern US, which exhibits the highest runoff ratio and the highest ratio of snow days. This cluster also shows extreme values with little variability, but this time mainly for climatic characteristics (lowest T_{MAX} and aridity index, AI; low PSI

and highest N_P). It does also have the highest root zone depth. Cluster C5, on the other hand, has distributed patches in different parts of the study region. This cluster has the highest baseflow index values in the dataset (aside from C9), and shows little variability as well as the highest values for % Sand. At the same time, it shows considerable variability in climate (e.g., T_{MAX} and MAP) and landscape characteristics (% AG and RZD).

The discussion in this section supports the earlier statement that such an empirical analysis cannot be the endpoint for classification, but rather a step along the way. The focus on streamflow means that we are limited in the degree of detail regarding hydrologic function that we can extract from such an integrated measure. However, it also allows for the regionalization of the signatures used, and enables an extension of the classification scheme to ungauged basins (Yadav et al., 2007). The limited availability of detailed descriptors of geology (certainly for a dataset covering a large region) suggests that we are also limited with respect to understanding subsurface controls (see similar issues in Oudin et al., 2008). And finally, the variability and environmental gradients in the dataset define what controls could even occur. There is of course no guarantee that different datasets, with different gradients, would not show other relationships between signatures and climate/landscape; or that these relationships would not change with the scale of analysis (Kennard et al., 2010). Therefore it is in the physical interpretation where the potential for generalization lies, rather than in the actual empirical result, and therefore more model/theory based analyses such as that of Carrillo et al. (2011) can be very useful as a supplement to the empirical studies.

6   Conclusions

The lack of a generally accepted catchment classification framework brought the question of what defines hydrologic similarity to the forefront of hydrologic science. Wagener et al. (2007) suggested that a classification framework, which is both descriptive and predictive, can be derived if it is based on the notion of catchment function and contains an explicit mapping between function (as observed in so-called signatures), climate and landscape characteristics. Here we provide a first test of this proposition in an empirical study utilizing 280 catchments across the Eastern US. This work provides insight into hydrologic similarity of catchments in the Eastern United States, and offers suggestions for controls of their hydrologic behavior.

We defined six signatures that can be derived from precipitation-temperature-streamflow data and used a Bayesian clustering algorithm to identify groups of similar catchments. Nine clusters with a relatively clear separation were identified. Spatially, most of the clusters exhibited some degree of connectivity suggesting that spatial proximity is a good indicator of similarity. It is likely that this

result is due to climatic and some landscape characteristics changing slowly in space. Further, the results suggest that permeable soils provide a buffer to how strongly a catchment responds to variability in climate. Our result therefore suggests that soil properties will modify the impacts of climate change on hydrologic regimes, which means that changes in precipitation and temperature will not impact the streamflow response equally. Assessing the implications of climate variability and change on hydrologic similarity will be the content of future research. Overall, the physical interpretation of why the members of a particular class behave similarly is very encouraging and demonstrates the merit of this kind of clustering analysis for understanding hydrological similarity and its causes. Expanding this kind of study using much larger, even global, datasets has the potential to provide further insight into catchment similarity, and, in combination with numerical modeling, can result in a general catchment classification framework.

Limitations of the study presented here are its purely empirical nature and the focus on streamflow as the only hydrologic response variable. However, signatures such as flow duration curves have been used for many years to define the hydrologic character of catchments and hence provide an excellent starting point for catchment classification. We further believe that the limitations of empirical studies can be aided by numerical experiments in which idealized systems are tested using catchment models. See the companion paper by Carrillo et al. (2011) for further an example application of this type.

Supplementary material related to this article is available online at: http://www.hydrol-earth-syst-sci.net/15/2895/2011/ hess-15-2895-2011-supplement.pdf.

References

Archcar, F., Camadro, J., and Mertivier, D.: AutoClass@IJM: a powerful tool for Bayesian classification of heterogeneous data in biology, Nucleic Acids Res., 37, 1–5, doi:10.1093/nar/gkp430, 2009.

Arnold, J. G. and Allen, P. M.: Automated methods for estimating baseflow and ground water recharge from streamflow records, J. Am. Water Resour. Ass., 35(2), 411–424, 1999.

Arnold, J. G., Allen, P. M., Muttiah, R., and Bernhardt, G.: Automated base flow separation and recession analysis techniques, Ground Water, 33, 1010–1018, 1995.

Bashford, K., Beven, K., and Young, P.: Observational data and scale- dependent parameterizations: Explorations using a virtual hydrological reality, Hydrol. Process., 16, 293–312, 2002.

Beven, K. J.: Uniqueness of place and process representations in hydrological modelling, Hydrol. Earth Syst. Sci., 4, 203–213, doi:10.5194/hess-4-203-2000, 2000.

Black, P. E.: Watershed functions, J. Am. Water Resour. Ass., 33, 1–11, 1997.

Bormann, H.: Towards a hydrologically motivated soil texture classification, Geoderma, 157, 142–153, 2010.

Bormann, H., Diekkrüger, B., and Renschler, C.: Regionalization concept for hydrological modelling on different scales using a physically based model: results andevaluation, Phys. Chem. Earth B, 24(7), 799–804, 1999.

Bower, D., Hannah, D. M., and McGregor, G. R.: Techniques for assessing the climatic sensitivity of river flow regimes, Hydrol. Process., 18, 2515–2543, 2004.

Broxton, P. D., Troch, P. A., and Lyon, S. W.: On the role of aspect to quantify water transit times in small mountainous catchments, Water Resour. Res., 45, W08427, doi:10.1029/2008WR007438, 2009.

Buttle, J.: Mapping first-order controls on streamflow from drainage basins: the T^3 template, Hydrol. Process., 20, 3415–3422, 2006.

Carrillo, G., Troch, P. A., Sivapalan, M., Wagener, T., Harman, C., and Sawicz, K.: Catchment classification: hydrological analysis of catchment behavior through process-based modeling along a climate gradient, Hydrol. Earth Syst. Sci. Discuss., 8, 4583–4640, doi:10.5194/hessd-8-4583-2011, 2011.

Cheeseman, P. and Stutz, J.: Bayesian Classification (AutoClass): Theory and Results, in: Advances in Knowledge Discovery and Data Mining, edited by: Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P., and Uthurusamy, R., AAAI Press/MIT Press, Cambridge, 1996.

Dooge, J. C. I.: Looking for hydrologic laws, Water Resour. Res., 22(9), 46–58, 1986.

Dooge, J. C. I.: Sensitivity of runoff to climate change: A Hortonian approach, B. Am. Meteorol. Soc., 73(12), 2013–2024, 1992.

Duan, Q., Schaake, J., Andreassian, V., Franks, S., Gupta, H. V., Gusev, Y. M., Habets, F., Hall, A., Hay, L., Hogue, T. S., Huang, M., Leavesley, G., Liang, X., Nasonova, O. N., Noilhan, J., Oudin, L., Sorooshian, S., Wagener, T., and Wood, E. F.: Model Parameter Estimation Experiment (MOPEX): Overview and Summary of the Second and Third Workshop Results, J. Hydrol., 320, 3–17, 2006.

Eckhardt, K.: How to construct recursive digital filters for Baseflow separation, Hydrol. Process., 19, 507–515, 2005.

Eckhardt, K.: A comparison of baseflow indices, which were calculated with seven different baseflow separation methods, J. Hydrol., 352, 168–173, 2007.

Gottschalk, L.: Hydrological regionalization of Sweden, Hydrol. Sci. J., 30(1), 65–83, 1985.

Grigg, D. B.: The logic of regional systems, Ann. Ass. Am. Geog., 55, 465–491, 1965.

Grigg, D. B.: Regions, Models and Classes, In: Models in Geography, Methuen, London, 467–509, 1967.

Gupta, H. V., Wagener, T., and Liu, Y.: Reconciling theory with

www.hydrol-earth-syst-sci.net/15/2895/2011/

Hydrol. Earth Syst. Sci., 15, 2895–2911, 2011

observations: elements of a diagnostic approach to model evaluation, Hydrol. Process, 22, 3802–3813, 2008.

Haines, A. T., Finlayson, B. L., and McMahon, T. A.: A global classification of river regimes, App. Geog., 8, 255–272, 1988.

Haltas, I. and Kavvas, M. L.: Scale invariance and self-similarity in hydrologic processes in space and time, J. Hydrol. Engr., 16(1), 51–63, 2011.

Harte, J.: Towards a synthesis of the Newtonian and Darwinian worldviews, Physics Today, 55(10), 29–34, 2002.

Hubert, L. and Arabie, P.: Comparing partitions, J. Classification, 2(1), 193–218, 1985.

Institute of Hydrology: (UK), Low-flow studies, Report No. 1, Wallingford, Oxon, UK, 1980.

Jain, A. K., Murty., M. N., and Flynn., P. J.: Data Clustering: A review, ACM Computer Surveys, 31(3), 1999.

Jensco, K. G., McGlynn, B. L., Gooseff, M. N., Wondzell, S. M., Bencala, K. E., and Marshall, L. A.: Hydrologic connectivity between landscapes and streams: Transferring reach- and plot-scale understanding to the catchment scale, Water Resour. Res., 45, W04428, doi:10.1029/2008WR007225, 2009.

Jothityangkoon, C., Sivapalan, M., and Farmer, D. L.: Process controls of water balance variability in a large semi-arid catchment: downward approach to hydrological model development, J. Hydrol., 254, 174–198, 2001.

Kennard, M. J., Mackay, S. J., Pusey, B. J., Olden, J. D., and Marsh, N.: Quantifying uncertainty in estimation of hydrologic metrics, River Res. Applic., 26, 137–156, 2010.

Kroll, C., Luz, J., Allen, B., and Vogel, R. M.: Devoloping a watershed characteristics database to improve low streamflow prediction, J. Hydrol. Eng., 2, 116–125, 2004.

Kullback, S. and Leibler, R. A.: On Information and Sufficiency, Ann. Math. Stat., 22(1), 79–86, doi:10.1214/aoms/1177729694, 1951.

Laaha, G. and Blöschl, G.: A comparison of low flow regionalization methods – catchment grouping, J. Hydrol., 323, 193–214, 2006.

Lim, K. J., Engel, B. A., Tang, Z., Choi, J., Kim, K., Muthukrishnan, S., and Tripathy, D.: Automated web GIS based hydrograph analysis tool, WHAT, J. Am. Water Resour. Ass., 04133, 1407–1416, 2005.

Lyon, S. W. and Troch, P. A.: Development and application of a catchment similarity index for subsurface flow, Water Resour. Res., 46, W03511, doi:10.1029/2009WR008500, 2010.

McDonnell, J. J. and Woods, R. A.: On the need for catchment classification, J. Hydrol., 299, 2–3, 2004.

McGlynn, B. L., McDonnell, J. J., Seibert, J., and Stewart, M. K.: On the relationships between catchment scale and streamwater mean residence time, Hydrol. Process., 17, 175–181, 2003.

McGuire, K. J., McDonnell, J. J., McGlynn, B. L., Weiler, M., Kendall, C., Welker, J. M., and Seibert, J.: Watershed residence time and the role of topography, Water Resour. Res., 41, W05002, doi:10.1029/2004WR003657, 2005.

McIntyre, N., Lee, H., Wheater, H. S., Young, A., and Wagener, T.: Ensemble prediction of runoff in ungauged watersheds, Water Resour. Res., 41, W12434, doi:10.1029/2005WR004289, 2005.

McMillan, H. K., Clark, M. P., Bowden, W.B., Duncan, M., and Woods, R. A.: Hydrological field data from a modeller's perspective: Part 1. Diagnostic tests for model structure, Hydrol. Proc, 25, 511–522, doi:10.1002/hyp.7841, 2010.

McNeil, V. H., Cox, M. E., and Preda, M.: Assessment of chemical water types and their spatial variation using multi-stage cluster analysis, Queensland, Australia, J. Hydrol, 310, 181–200, 2005.

Merz, R. and Blöschl, G.: Regionalisation of catchment model parameters, J. Hydrol., 287, 95–123, 2004.

Merz, R. and Blöschl, G.: Flood frequency regionalisation – spatial proximity vs. catchment attributes, J. Hydrol., 302, 283–306, 2005.

Merz, R. and Blöschl, G.: A regional analysis of event runoff coefficients with respect to climate and catchment characteristics in Austria, Water Resour. Res., 45, W01405, doi:10.1029/2008WR007163, 2009.

Milly, P. C. D.: Climate, soil water storage, and the average water balance, Water Resour. Res., 30, 2143–2156, 1994.

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., and Stouffer, R. J.: Stationarity is dead: Whither water management?, Science, 319, 573–574, 2008.

Moliere, D. R., Lowry, J. B. C., and Humphrey, C. L.: Classifying the flow regime of data limited streams in the wet–dry tropical region of Australia, J. Hydrol, 367(12), 1–13, 2009.

Monk, W. A., Wood, P. J., Hannah, D. M., and Wilson, D. A.: Selection of river flow indices for the assessment of hydroecological change, River Res. Appl., 23, 113–122, 2007.

Morin, E., Georgakakos, K. P., Shamir, U., Garti, R., and Enzel, Y.: Objective, observations-based, automatic estimation of the catchment response timescale, Water Resour. Res., 38(10), 1212, doi:10.1029/2001WR000808, 2002.

Olden, J. D. and Poff, N. L.: Redundancy and the choice of hydrologic indices for characterizing streamflow regimes, River Res. App., 19, 101–121, 2003.

Omernik, J. M.: Ecoregions of the conterminous United States, Ann. Ass. Am. Geog., 77, 118–125, 1987.

Oudin, L., Andréassian, V., Perrin, C., Michel, C., and Le Moine, N.: Spatial proximity, physical similarity, regression and ungaged catchments: A comparison of regionalization approaches based on 913 French catchments, Water Resour. Res., 44, W03413, doi:10.1029/2007WR006240, 2008.

Oudin, L., Kay, A., Andreassian, V., and Perrin, C.: Are seemingly physically similar catchments truly hydrologically similar? Water Resour. Res., 46, W11558, 2010.

Panda, U. C., Sundaray, S. K., Rath, P., Nayak, B. B., and Bhatta, D.: Application of factor and cluster analysis for characterization of river and estuarine water systems – a case study: Mahanadi River (India), J. Hydrol., 331, 434–445, 2006.

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

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

Poff, N., Olden, J. D., Pepin, D. M., and Bledsoe, B. P.: Placing global stream flow variability in geographic and geomorphic contexts, River Res. App., 22, 149–166, 2006.

Pryor, S. C. and Schoof, J. T.: Changes in the seasonality of precipitation over the contiguous USA, J. Geophys. Res., 113, D21108, doi:10.1029/2008JD010251, 2008.

Ramachandra Rao, A. and Srinivas, V.V.: Regionalization of water-

Hydrol. Earth Syst. Sci., 15, 2895–2911, 2011

www.hydrol-earth-syst-sci.net/15/2895/2011/

sheds by hybrid cluster analysis, J. Hydrol., 318, 37–56, 2006.

Rand, W. M.: Objective criteria for the evaluation of clustering methods, J. Am. Stat. Assoc., 66(336), 846–850, 1971.

Robinson, J. S., Sivapalan, M., and Snell., J. D.: On the relative roles of hillslope processes, channel routing, and network geomorphology in the hydrologic response of natural catchments, Water Resour. Res., 31, 3089–3101, 1995.

Rosero, E., Yang, Z.-L., Wagener, T., Gulden, L. E., Yatheendradas, S., and Niu, G.-Y.: Quantifying parameter sensitivity, interaction and transferability in hydrologically enhanced versions of the Noah-LSM over transition zones during the warm season, J. Geophys. Res., 115, D03106, doi:10.1029/2009JD012035, 2010.

Samaniego, L., Bardossy, A., and Kumar, R.: Streamflow prediction in ungauged catchments using copula-based dissimilarity measures, Water Resour. Res., 46, W02506, doi:10.1029/2008WR007695, 2010.

Sankarasubramanian, A. and Vogel, R. M.: Hydroclimatology of the continental United States, Geo. Res. Lett., 30, 1363, doi:10.1029/2002GL015937, 2003.

Sankarasubramanian, A., Vogel, R. M., and Limbrunner, J. F.: Climate elasticity of streamflow in the United States, Water Resour. Res., 377, 1771–1781, 2001.

Santhi, C., Allen, P. M., Muttiah, R. S., Arnold, J. G., and Tuppad, P.: Regional estimation of base flow for the conterminous United States by hydrologic landscape regions, J. Hydrol., 351, 139–153, 2008.

Schaake, J. C.: From Slimate to flow, in: Climate Change and US Water Resources, edited by: Waggoner, P. E., 8, 177–206, John Wiley, New York, 1990.

Schaake, J. C., Duan, Q., Smith, M., and Koren, V.: Criteria to select basins for hydrologic model development and testing, Preprints, 15th Conf. On Hydrology, Long Beach, CA, Am. Meteorol. Soc., 10–14 January 2000, Paper P1.8. 2000.

Schaake, J., Cong. C., and Duan, Q.: The US MOPEX Data Set, Large Sample BasinExperiments for Hydrological Model Parameterization, 9–26, IAHS Publ. 307, 2006.

Shamir, E., Imam, B., Gupta, H. V., and Sorooshian, S.: Application of temporal streamflow descriptors in hydrologic model parameter estimation, Water Resour. Res., 41, W06021, doi:10.1029/2004WR003409, 2005.

Sivapalan, M.: Pattern, process and function: Elements of a unified theory of hydrology at the catchment scale, in: Encyclopedia of Hydrological Sciences, edited by: Anderson, M., London, John Wiley, 193–219, 2005.

Stutz, J. and Cheeseman, P.: AutoClass – a Bayesian Approach to Classification, Maximum Entropy and Bayesian Methods, Cambridge 1994, edited by: Skilling, J. and Sibisi, S., Kluwer Academic Publishers, Dordrecht, 1995.

Tallaksen, L. M. and van Lanen, H. A. J. (Eds.): Hydrological drought, Elsevier, Amsterdam, NL, 2004.

Tetzlaff, D., Seibert, J., and Soulsby, C.: Inter-catchment comparison to assess the influence of topography and soils on catchment transit times in a geomorphic province; the Cairngorm mountains, Scotland, Hydrol. Process., 23, 1874–1886, 2009.

van Werkhoven, K., Wagener, T., Reed, P., and Tang, Y.: Rainfall characteristics define the value of streamflow observations for distributed watershed model identification, Geophys. Res. Lett., 35, L11403, doi:10.1029/2008GL034162, 2008.

Vogel, R. M. and Fennessey, N. M.: Flow Duration Curves I: A New Interpretation and Confidence Intervals, ASCE J. Water Resour. Plan. Managem., 120, 485–504, doi:10.1061/(ASCE)0733-9496, 1994.

Vogel, R. M. and Kroll, C. N.: Regional geohydrologic-geomorphic relationships for the estimation of low-flow statistics, Water Resour. Res., 28, 2451–2458, 1992.

Wagener, T., Sivapalan, M., Troch, P. A., and Woods, R. A.: Catchment Classification and Hydrologic Similarity, Geog. Comp., 1/4, 901–931, 2007.

Wagener, T., Sivapalan, M., and McGlynn, B. L.: Catchment classification and services – Toward a new paradigm for catchment hydrology driven by societal needs, in: Encyclopedia of Hydrological Sciences, edited by: Anderson, M.G., John Wiley & Sons Ltd, 2008.

Wagener, T., Sivapalan, M., Troch, P. A., McGlynn, B. L., Harman, Gupta, C. J., Kumar, H. V., Rao, P. S. C., Basu, N. B., and Wilson, J. S.: The future of hydrology: An evolving science for a changing world, Water Resour. Res., 46, W05301, doi:10.1029/2009WR008906, 2010.

Webb, J. A., Bond, N. R., Wealands, S. R., McNally, R., Quinn, G. P., Vesk, P. A., and Grace, M. R.: Bayesian clustering with AutoClass explicitly recognizes uncertainties in landscape classification, Ecography, 30, 526–536, 2007.

Weiler, M., McGlynn, B. L., McGuire, K. J., and McDonnell, J. J.: How does rainfall become runoff? A combined tracer and runoff transfer function approach, Water Resour. Res., 39(11), 1315, doi:10.1029/2003WR002331, 2003.

Weiler, M. and McDonnell, J.: Virtual experiments: A new approach for improving process conceptualization in hillslope hydrology, J. Hydrol., 285(1–4), 3–18, 2004.

Winter, C. L., Springer, E. P., Costigan, K., Fasel, P., Mniszewski, S., and Zyvoloski, G.: Virtual watersheds: Simulating the water balance of the Rio Grande basin, Comput. Sci. Eng., 6(3), 18–26, 2004.

Winter, T. C.: The concept of hydrologic landscapes, J. Am. Water Resour. Ass., 37, 335–349, 2001.

Wittenberg, H. and Sivapalan, M.: Watershed groundwater balance estimation using streamflow recession analysis and baseflow separation, J. Hydrol., 219, 20–33, 1999.

Wolock, D. M., Winter, T. C., and McMahon, G.: Delineation and evaluation of hydrologic-landscape regions in the United States using geographic information system tools and multivariate statistical analyses, Environ. Manage., 34, 71–88, 2004.

Woods, R. A.: The relative roles of climate, soil, vegetation and topography in determining seasonal and long-term catchment dynamics, Adv. Water Resour., 26(3), 295–309, 2003.

Yadav, M., Wagener, T., and Gupta, H.: Regionalization of constraints on expected watershed response, Adv. Water Resour., 30, 1756–1774, 2007.

Zhang, Z., Wagener, T., Reed, P., and Bhushan, R.: Reducing uncertainty in predictions in ungauged basins by combining hydrologic indices regionalization and multiobjective optimization, Water Resour. Res., 44, W00B04, doi:10.1029/2008WR006833, 2008.