The Pennsylvania State University

The Graduate School

CATCHMENT CLASSIFICATION: AN EMPIRICAL ANALYSIS OF HYDROLOGIC SIMILARITY

A Thesis in

Civil Engineering

by

Keith Alexander Sawicz

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

August 2009

The thesis of Keith Alexander Sawicz was reviewed and approved* by the following:

Thorsten Wagener
Associate Professor of Civil and Environmental Engineering

Michael N. Gooseff
Assistant Professor of Civil and Environmental Engineering

Peggy Johnson
Professor of Civil and Environmental Engineering
Head of the Department of Department or Graduate Program

*Signatures are on file in the Graduate School

ABSTRACT

Classification is a fundamental process of understanding major attributes and interactions of those attributes in any field of science. Hydrology has lacked such a classification system due to a lack of understanding of the dominant hydrologic functions. This study offers a robust empirical method of classifying the basic units of hydrology (catchments) into groups based on hydrologic similarity. Approximately 300 catchments were chosen for this study located east of the Rocky Mountains extending out towards the Atlantic Ocean. Correlation and distribution tests for over 40 signatures was completed, and the six signatures showing the high amounts of information (R_{QP}, S_{FDC}, I_{BF}, E_{QP},R_{SD} and b) were analyzed and grouped using a hierarchical cluster analysis. Results were displayed through the use of dendrograms, recursive pattern plots, and spatial distribution plots to extract information from the individual signatures and groups at many levels. A classification system is suggested based on the signatures selected and an explanation of the possible hydrologic processes occurring in each cluster is offered. Applications of this study include but are not limited to improvements in hydrologic process understanding, a parameter independent method of confining hydrologic model outputs, and the ability to predict the impact of changes experienced at a particular catchment from human impacts or climate nonstationarity.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Acknowledgements

I owe a great deal of thanks to many people who have helped and influenced me throughout graduate school. I would like to thank Dr. Thorsten Wagener for enabling and guiding me through this project. Through our many interactions, whether in his office, at one of the many conferences we attended, or at a local coffee shop, I have strengthened my weaknesses and let my potential be known. He has acted as a wonderful advisor, mentor, and friend. I would like to thank the other members of my committee, Dr. Michael Gooseff and Dr. Peggy Johnson for the comments and discussions this work. I would like to thank the other members of my research project including Gustavo Carrillo and Dr. Peter Troch from the University of Arizona, and Siva Sivapalan of the University of Illinois, Urbana-Champaign, without whom this project would not be possible. Our constant collaboration through teleconferences and constant streams of email resulted in the creation of project parameters through use a true depth of knowledge that they bring to the project.

I would like to thank the other members of the Thorsten Wagener group including Katie van Werkhoven and Christa Kelleher and others for their guidance and support. I would also like to thank Joseph Kasprzyk, Joshua Kollat, Adam Ward, Corbitt Kerr, George Holmes III, and many many other friends for making graduate school life a continuously stimulating and enjoyable experience. I would like to give special thanks to Mukesh Kumar with regard to the many late night hours we have spent discussing the problems and nature of the world, as well as offering me continual support and guidance. I would also like to give special thanks to my family, which whom I would not be the person I am today. In such a non-stationary world we live in, they have helped me to stay grounded, but always reminded me to look towards the stars.

Dedication

To my grandparents, Myra and Michael, to my parents, Robert and Claire, and to my siblings, Jason, Michael, and Stephanie

Chapter 1

Introduction

1.1 Motivation

Catchments are the topographically defined spatial units that theoretically collect all precipitation and channel it to a common outlet, assuming no evaporative losses. This spatial unit represents a basic entity commonly used in hydrology. The catchment forms a landscape element (at various scales) that integrates all aspects of the hydrological cycle within a defined area that can be studied, quantified, and acted upon (Wagener et al., 2007). Catchments vary widely in physical characteristics (ie. vegetation, landscape, soils, ecology, ect.) and climatic characteristics (ie. precipitation, temperature, energy). Despite the uniqueness and complexity that each catchment exhibits (Beven, 2000), they generally also show some level of organization, which suggests that their functional behavior should be predictable (Sivapalan, 2005).

Webster's English dictionary defines classification as a *systematic arrangement in groups or categories according to established criteria*. Classification of the central entity of interest lies at the heart of many sciences. The classifications of organisms in biology or of elements in chemistry are well-known examples. An important task of science in any field lies in continuously organizing the body of knowledge gained by scientific inquiry. The science of 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 (see detailed discussions in McDonnell and Woods 2004 and in Wagener et al. 2007).

Identifying and categorizing dominant catchment functions in relation to hydrologic variables such as streamflow would quantify the degree of similarity that is exists between catchments (McDonnell and Woods 2004; Wagener et al. 2007).

The goals of a widely accepted catchment classification system for hydrology would (see McDonnell and Woods, 2004; Wagener et al., 2007):

- *provide an important organizing principle in itself, complementing the concept of the hydrological cycle and the principle of mass conservation;*
- *help with both modeling and experimental approaches to hydrology, by providing guidance on the similarities and differences between catchments;*
- *improve communication by providing a common language for discussions;*
- *allow the rational testing of hypotheses about the similarity/dissimilarity of hydrological systems from around the globe, as well as better design of experimental and monitoring networks by focusing on measuring the most important controls;*
- *provide better guidance for choosing appropriate models for poorly understood hydrological systems;*
- *be a major advancement toward guidance for the applicability of various simulation methods for predictions in ungauged basins;*
- *provide constraints and diagnostic metrics that can be used for model evaluation/diagnostics and application and ungauged locations; and*
- *provide, to first order, insights into the potential impacts of land use and climate changes on the catchment-scale hydrologic response in different parts of the world*.

An additional issue that should be considered in any catchment classification system is the nonstationarity of physical climatic characteristics (Milly et al., 2008). A good classification system will only be robust if it inherently represents enough information to describe change or evolution of catchments from one group classification to another. This idea is similar to the idea of an organism evolving over time. The biological classification system represents enough organism attributes to reclassify this new organism without requiring the classification system itself to change. Land use and climate change and evolution will alter physical catchment

characteristics and therefore the hydrologic response is also likely to change. Any long-term classification system has to be able to account for this evolution and be dynamic in nature.

1.2 Review

Past approaches to catchment classification have been based on grouping physically and/or climatically similar catchments or catchments that behave similar with respect to some hydrological variable such as streamflow. Winter (2001) defines catchments to be similar with respect to climate, topography and geology as hydrologic landscapes, assuming that catchments are similar with respect to these three criteria will also behave similar in a hydrological sense (Winter, 2001). This idea was applied by Beighley et al. (2005) uses impervious land cover, ground slope, and depth to bedrock to describe hydrologic spatial units. However, in the Beighley et al. (2005) study, the selection of values to relate physical values to hydrologic signatures are arbitrary and fail to include climate characteristics due to a small sample size of catchments within close proximity of one another.

Assessing similarity in terms of river regime characteristics is widely used in ecology. For example, Haines et al. (1988) classified river regimes in terms of seasonality of flow that the rivers in this study experience. Merz and Bloschl (2005) and Laaha and Bloschl (2006) tested the value of similarity and of proximity in the context of understanding flood and low flow characteristics respectively. In studies conducted thus far that focused on grouping catchments similar with respect to their response, there has been a lack in understanding as to which hydrologic signatures should be used to classify and characterize a catchment. In order to properly identify key hydrologic characteristics used to differentiate hydrologic landscapes, there is a need in hydrology for a catchment classification system based on hydrologic similarity and process understanding (McDonnell and Woods 2004). But what does hydrologic similarity mean?

Wagener et al. (2007) first offered a method of dissecting a catchment into its hydrologically defined functions through which water moves through a catchment. These functions include partitioning, storage, and release functions; and connected those to observable catchment signatures. Wagener et al. (2008) later expanded this idea by incorporating services that are provided by the catchment due to the provision of freshwater.

1.3 Classification based on hydrologic similarity

Ultimately, our interest lies in the creation of a catchment classification system that groups catchments with similar hydrologic behavior defined by similarity in hydrologic functions. Hydrologic functions are the main physical components of how a catchment responds to precipitation events (e.g. streamflow, groundwater, soil moisture). The fashion in which water eventually leaves the catchment can be defined as the release function. If the release functions of two or more catchments are similar, it can be stated that these catchments are hydrologically similar. We assume that this similarity can be captured by a few carefully selected release characteristics of the catchment, defined as key hydrologic signatures. If key signatures and the functions they define can be predicted from observable physical and climatic catchment characteristics, then we would have a generally applicable classification system. Such a classification can change when any of the predictive variables change.

Many other strategies to catchment classification begin with the opposite strategy by basing their grouping of catchments on similarity in physical characteristics (e.g. Merz and Bloschl, 2005). They then assume that catchments with similar physical characteristics behave hydrologically similar if located in similar climatic regions. This approach requires strong assumptions about how physical and signatures relate. In this study, we will approach the problem in an empirical way to avoid making such a priori assumptions.

1.4 Objectives and Scope

The main objectives of this study are to: [1] identify key signatures derived from the time series information including streamflow, precipitation, and temperature, and [2] to define and distinguish controlling hydrologic functions between catchments formalized in a classification system. Deriving and understanding hydrologically representative signatures is imperative for a robust and informational classification system. Key signatures that contain sufficient information are needed to understand and capture the functional behavior of a catchment in order to group catchments of similar key signature values and therefore dominant catchment functions.

The study uses about 300 US catchments located in the Eastern half of the US. These catchments represent a wide range in signature values in each of the key signatures chosen, along with a wide variety of physical and climatic characteristics. Focusing on this area of the US allows the focus to be on areas that have relatively high stream gauge density. Stream gauge density decreases sharply in catchments located west of -103^{o} longitude. The catchment release options considered are streamflow and evapotranspiration under the assumption that net groundwater fluxes are negligible over longer periods of time. To calculate annual and inter-annual signatures, the hydrologic year was used, beginning on October 1st and ending on September 30th of each year. This was done in order to lessen the effect of snow storage between years. Streamflow, precipitation, and temperature data is available for a large number of catchments in the form of a daily time series, and was therefore used in this study to derive hydrologic signatures.

Chapter 2

Background

2.1 Conceptual Model of Catchment Function

Catchment functions are defined here as the physical processes that partition, store, and release water through the catchment (Black, 1997; Wagener et al., 2007). Figure 2-1 shows the conceptual model of catchment function and catchment signatures used in this study as a basis for a new catchment classification system. This conceptual model was modified from Wagener et al. (2007) and Wagener et al. (2008), incorporating human impacts to catchments.

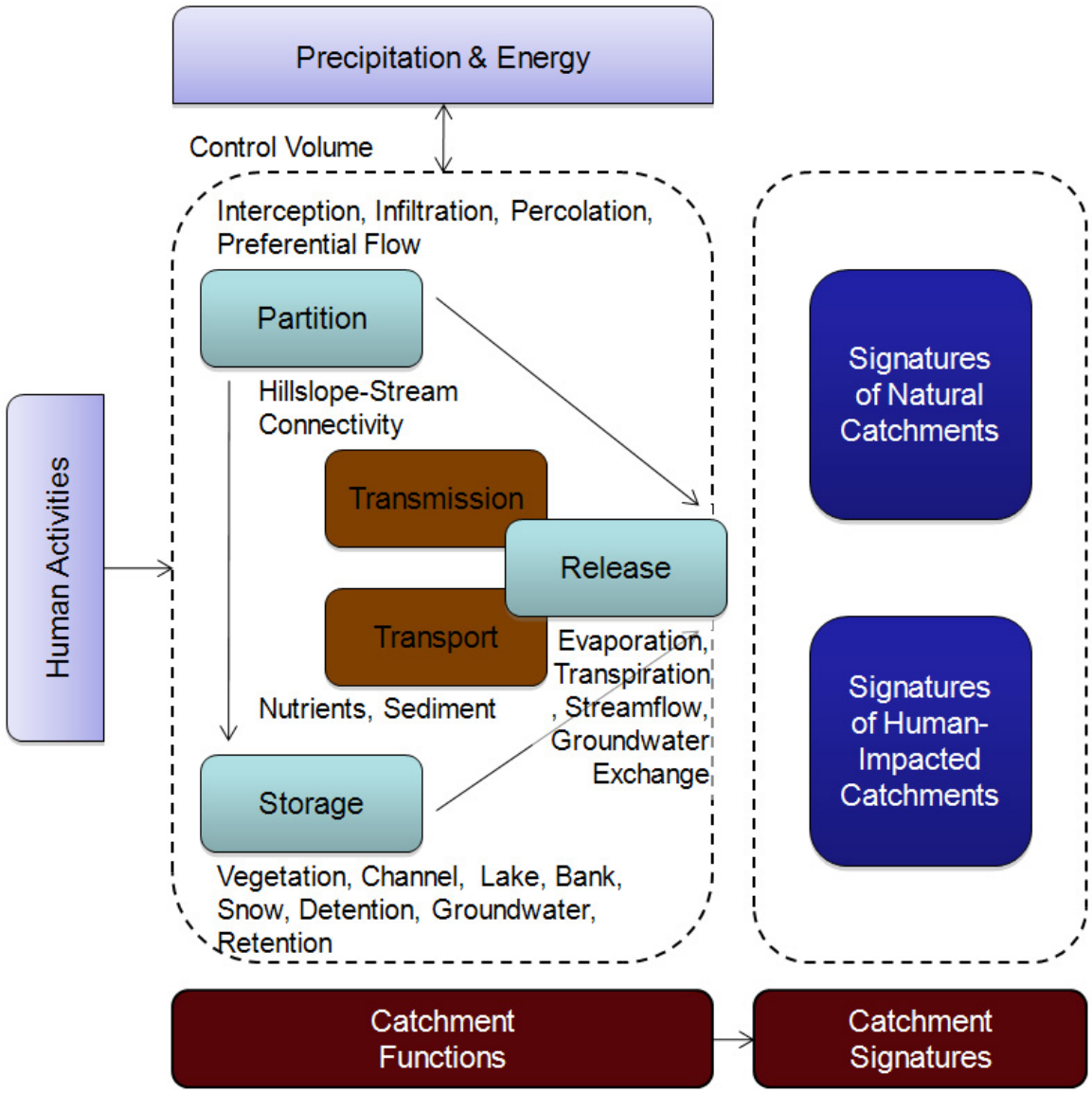


Figure 2-1. Conceptual model of catchment function and catchment signatures. Human activities impact catchment functions, which result in a release function that can be dissected into signatures of natural and human impacted catchments.

This conceptual model forms the basis of the proposed hydrologic classification framework. The catchment is treated as a control volume in which the catchment acts as a filter to its inputs. Inputs of this model include precipitation and energy. Partitioning is the process of dividing precipitation into its components as the water is being introduced into the catchment.

This process begins above the ground surface as water is intercepted by leaf cover and vegetation. Any water that falls through this cover reaches the ground surface, where it is further partitioned into soil moisture, groundwater, or surface water flow. Once the water that enters the catchment is partitioned, it is then conveyed to storage and release functions.

Storage of precipitation in the catchment refers to how long and where partitioned rainfall is being stored before its eventual release out of the catchment. Types of storage within a catchment include soil moisture (unsaturated zone), groundwater (saturated zone), surface storage (lakes, wetlands), above surface (leaf interception zone), and vegetation types. The retention times of each of these storage types vary greatly throughout the catchment. As an example, water storage retention time on leaf interception is much shorter than water storage within groundwater. Another type of storage that may occur based on climatic conditions (<2.^{o}C) is snow, ice cover, or permafrost. The predominate type of storage the catchment is based on the particular climate and catchment characteristics of the location under study. Once the storage function is complete, water enters the release function.

Release of stored water is defined as how the water ultimately leaves the catchment area. This release can generally be characterized as evapotranspiration, surface runoff, or groundwater flow (excluding human abstractions). The governing mechanisms to differentiate between types of release functions are the dominant storage function and climate forcing. If a large portion of storage in a catchment is experience on leaf cover, in the unsaturated area of high vegetation density, or in an area with excess energy coming into the catchment, it is likely that evapotranspiration is a controlling factor of this function. If energy introduced into the system is low and vegetation is sparse, then it is likely that most of the water within the catchment will flow along the surface or through the groundwater, exiting the catchment as streamflow. The catchment response defined by its functions should be observable in the catchment signatures. A signature can be defined as a characteristic of a dynamic hydrologic variable (e.g. streamflow) that provides insight into the catchment function and that is based on hydrologic theory

2.2 Catchment Signatures

In the context of this paper, we will derive all signatures from streamflow, precipitation, and temperature observations due to the availability of long term observations of this variable for all catchments. In general, signatures can also be derived from other variables such soil moisture,

groundwater etc. The time-series of streamflow exiting the catchment is represented by a hydrograph. These hydrographs can be broken down and characterized into separate pieces of information known as streamflow signatures. Signatures represent information about the types and degrees of functions occurring within a catchment. Key signatures (signatures that convey information independent of other signatures and are considered uncorrelated) contain information that can identify a function or a collection of dependent functions. The question is: What is the minimum number of signatures needed to separate catchments with different functional behavior?

Signatures have been shown to be useful in constraining hydrologic model output (continuous streamflow predictions) for predictions in ungauged basins studies (Yadav et al., 2007). The study by Yadav et al. (2007) introduced a model independent approach as alternative to parameter regionalization and to a priori parameter estimation. Main advantage of the signature regionalization approach is that the relationship between signatures and physical/climatic catchment characteristics is less noisy than with the direct parameter estimation approaches. Zhang et al. (2008) frame the search for behavioral models within the framework of Yadav et al. (2007) in a multi-objective optimization framework to increase the efficiency of identifying feasible parameter sets. Bulygina et al. (2009) further expanded this approach in a Bayesian framework to derive statistically correct prediction limits.

Chapter 3

Methods

3.1 Streamflow Signatures used for Catchment Classification

To define key signatures in terms of physical function and processes, a conceptual representation of a catchment is needed. A proper physically based conceptual model should represent dominant real world processes while at the same time simplifying reality to something that is applicable to all catchments studied and that can convey understanding. The conceptual model that was used in this study can be seen in Figure 3-1, and a listing the parameters used and brief explanations of those parameters can be found in Table 3-1. These parameters aid in the explanation and definition of physical properties and their relationship to hydrologic functions.

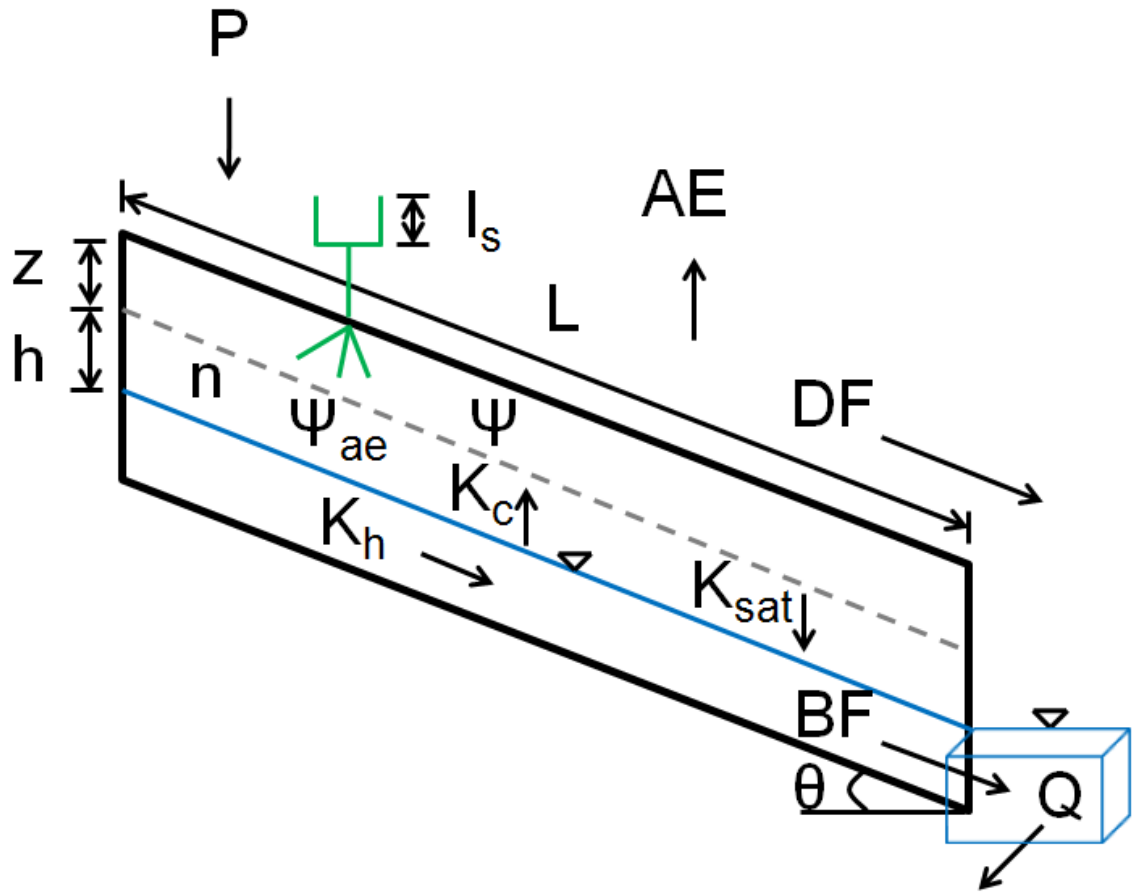


Figure 3-1. Schematic of catchment conceptualization. The blue line represents the groundwater level and the blue box represents the stream. This figure introduces the variables and parameters used to describe catchment physical characteristics and hydrologic fluxes explained in detail in Table 1.

Table 3-1. Name, units, name, and description of parameters used in the catchment conceptualization.

Name	Units	Full name	Description
AE	[L]	Actual evapotranspiration	Amount of water leaving catchment through evapotranspiration
BF	[-]	Base flow	Amount of water from groundwater table contributing to streamflow
DF	[-]	Direct flow	Amount of water from surface flow contributing to

			streamflow
Is	[L]	Interception storage	Amount of water intercepted and stored on leaf surfaces
Kh	[L/T]	Horizontal hydraulic conductivity	Average velocity of groundwater parallel to the bedrock.
Ksat	[L/T]	Saturated hydraulic conductivity	Hydraulic conductivity at soil saturation point
Kc	[L/T]	Vertical upwards hydraulic conductivity	Water movement in soil due to capillary forces
L	[L]	Length to stream	Average distance of a hillslope within a catchment.
N	[-]	Porosity	Fraction of soil made up of air or water
P	[L]	Mean annual precipitation	Long term measure of precipitation depth
Q	[L]	Mean annual streamflow	Long term measure of streamflow depth
Z	[L]	Root zone depth	Average depth of root zone dependant on the type of vegetation
Ψ	[-]	FC-WP or FC-Residual Moisture	Fraction of water volume to total soil volume found within the soil available to plants
ae	[L]	Capillary fringe depth	Length of capillary moisture depth based on soil type
H	[L]	Depth to Groundwater	Distance between the bottom of the root zone depth and the top of the ground water level.
Θ	[-]	Catchment Slope	Average slope found on the catchment

A total of 45 signatures were calculated for each of the 280 catchments based on streamflow, precipitation, and temperature information. Due to the large number of catchments included in this study, signature calculations need to be automated (as opposed to manual calculation). Automated calculation of signatures allows for [1] a universal method of calculating signatures and [2] an efficient way of calculating signatures from a large dataset. The distribution and range of signatures was also considered to ensure that values would be fully represented. A summary of all signatures calculated can be found in table A-3. From this list of signatures, key

signatures were identified and chosen for the remainder of the analysis. Key signatures were chosen from inspection using both quantifiable measures of correlation between signatures and understanding what the signatures represent in a hydrologic sense. The signatures considered to contain the most amount of information are physically represented in Figure 3-2. The signatures include Runoff Ratio, Base Flow Index, Slope of the Flow Duration Curve, Streamflow Elasticity, Recession Coefficient

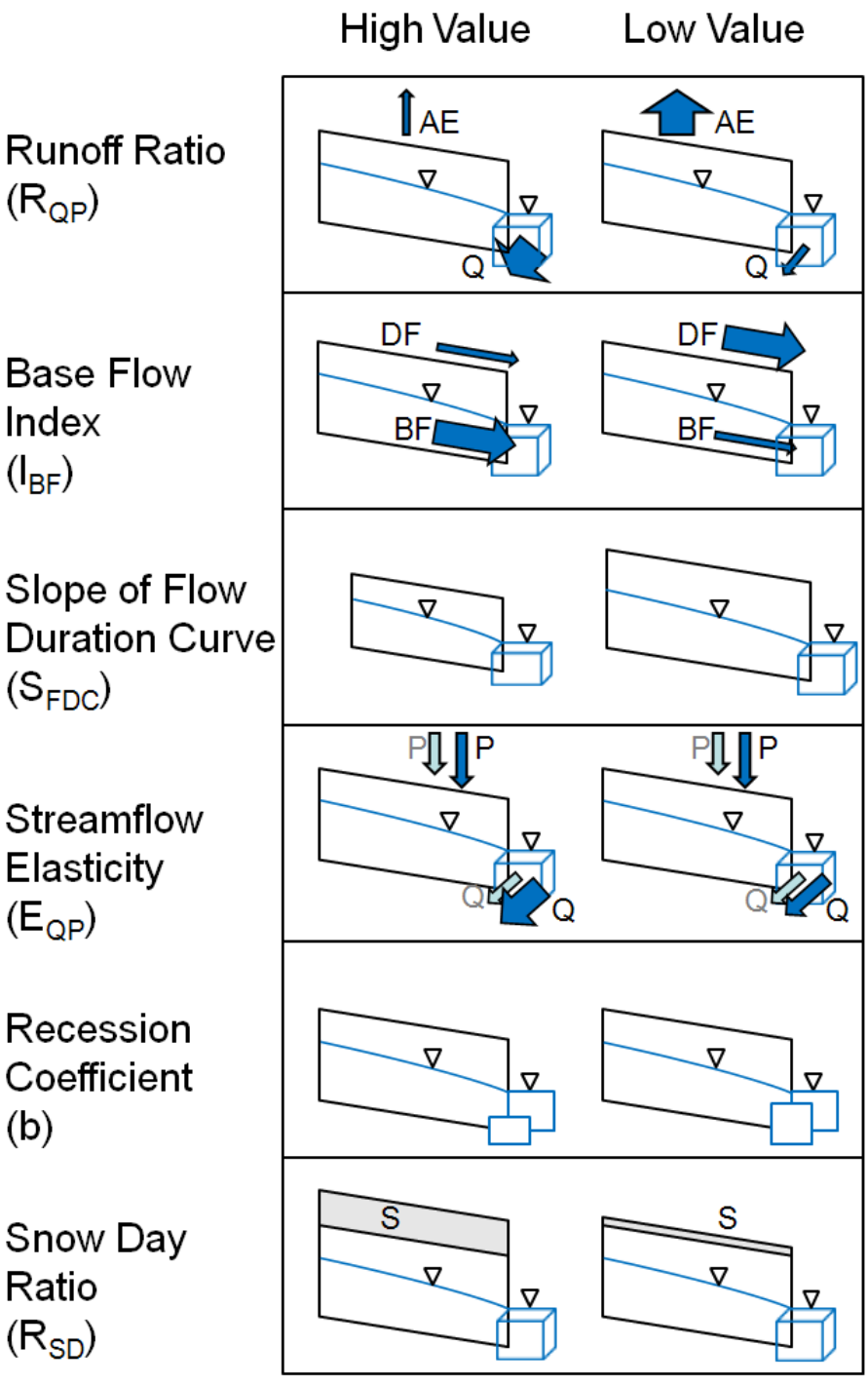


Figure 3-2. Conceptual representation of the six streamflow-based signatures suitable to define the functional behavior of catchments demonstrated as high and low values of each signature. Demonstrates in a physical sense what is meant by high and low values. The variable "S" stands for the amount of snow storage found within the catchment.

A particular value of a key signature is the explanatory power it holds and the ability to differentiate which hydrologic functions are dominant within the catchment. Using the conceptual model as a basis to display these signatures, Figure 3-2 shows the physical meaning of a high or low value of each key signature.

Runoff Ratio.

Runoff Ratio (R_{QP} [-]) is the ratio of long-term average streamflow to long-term average precipitation. A water balance characteristic such as R_{QP} is a very important signature in understanding how much water is available and how it leaves the catchment (Milly, 1994; Olden and Poff, 2003; Poff et al., 2003, Sankarasubramanian et al., 2001; Yadav et al., 2007). It represents the long term water balance separation between water exiting as streamflow or as evapotranspiration (assuming no net change in groundwater storage over a long time). A high runoff ratio identifies a large amount of water exiting the catchment as streamflow whereas a low value of runoff ratio identifies a large amount of water exiting the catchment as evapotranspiration. A physical conceptualization can be seen in Figure 3-2.

$$R_{QP} = \frac{Q}{P} \qquad \text{Eq 1.}$$

Where Q is the mean annual streamflow [mm], and P is the mean annual precipitation [mm].

Baseflow Index

Baseflow index (I_{BF} [-]) is the long term ratio of baseflow by total streamflow and is the summation of values of baseflow and streamflow calculated at the daily time step (Arnold et al.,1995, Vogel and Kroll, 1992; Kroll et al., 2004). A high value of I_{BF} means that a large portion of the streamflow is comprised of slow responding groundwater contributions, and a low value of I_{BF} means that a large portion of the streamflow is comprised of fast responding surface and interflow contributions. This index is most often derived by separating quick flow and baseflow using an automated or manual process. The manual process is time consuming and inappropriate for an empirical study of this scale. A range of automated methods have been proposed to

perform this separation (Kroll et al., 2004; Eckhardt, 2005; Institute of Hydrology, 1980; Arnold et al., 1995, Arnold and Allen, 1999). We use the one-parameterone pass digital filter method (DFM) based on the study by Arnold et al. (1995) and Lim et al. (2005).

Studies have shown that I_{BF} calculated using automatic separation techniques are highly correlated with one another, and the DFM has been show as a good representation of an automatic separation technique (Arnold et al., 1995; Lim et al., 2005). We do not consider the specific choice of method crucial since we are mainly interested in the relative differences between catchments, rather than in actual values.

$$Q_{D\,t} = a\,Q_{D\,t-1} + \frac{1+a}{2}\,(Q_t - Q_{t-1}) \qquad \text{Eq 2.}$$

Where $Q_{D\,n}$ is the direct flow value at time-step t, Q_t is the total flow at time step t, and a is a parameter coefficient. The value of α was set at a value of 0.925 based on a comprehensive case study performed by Eckhardt (2007). Eckhardt's study includes 65 USGS gauges that were randomly selected to compare seven different types of automated baseflow separation techniques. Values of I_{BF} calculated using a value of 0.925 for a DMF by Eckhardt showed a large range from 0.35 to 0.83 and found a correlation of over 0.92 when DFM is compared with the six other baseflow separation methods. Technical details of the study can be found in Eckhardt (2007).

$$Q_{B\,t} = Q_t - Q_{D\,t} \qquad \text{Eq 3.}$$

$Q_{B\,t}$ is the base flow value at time-step t. If the amount of direct flow or base flow is summated over the length of the time series, this would represent the total volume of direct flow or base flow, respectively.

$$I_{BF} = \sum \frac{Q_B}{Q} \qquad \text{Eq 4.}$$

Base flow is summated and normalized by the total amount of streamflow measured. The resulting dimensionless ratio is the Base Flow Index.

Slope of the Flow Duration Curve

Flow Duration Curve (FDC) is a frequency distribution formed from daily streamflow data of streamflow exceedance probability (Yadav et al., 2007; Jencso et al., 2009). The FDC is usually plotted on semi-log scale to better represent low flows. It represents a useful tool to analyze the variability in streamflow that a catchment exhibits.

To concentrate the information of variability in flow into a single value, a slope of the FDC is calculated. Slope of the FDC (S_{FDC} [-]) is calculated between the 33rd and 66th streamflow percentiles, because at semi-log scale this represents a relatively linear part of the FDC (Yadav et al., 2007; Zhang et al., 2008). A high slope values means flashy response and small storage/filter potential while a low slope value means slow response and large storage/filter potential.

$$S_{FDC} = \frac{log(Q_{33\%}) - \log(Q_{66\%})}{(66-33)} \quad \text{Eq 5.}$$

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

Inter-Annual Streamflow Elasticity

Inter-annual streamflow elasticity (E_{QP}[-]) calculated from the median of the series of the difference between annual streamflow divided by the difference between annual precipitation normalized by the long-term runoff ratio. A small value (less than one) would be defined as a percent change in precipitation resulting in a smaller percent change in streamflow between years. In other words, a value less than one would define the catchment's streamflow as relatively insensitive to precipitation. A value greater than one would mean that a percent change in precipitation between years would result in a larger percent change in streamflow. In other words, a value greater than on would define the catchment as sensitive to precipitation. A value of one would mean that there is a linear relationship between precipitation and streamflow. The median value is taken as a robust measure, since the median number in a dataset filters out the outliers (Sankarasubramanian et al., 2001; Sankarasubramanian and Vogel, 2003).

$$\varepsilon_{QP} = median(\frac{dQ}{dP}\frac{P}{Q}) \quad \text{Eq 6.}$$

Where ε_{QP} is the streamflow elasticity, dQ is the difference between the previous year's streamflow and the current year's streamflow, dP is the difference between the previous year's streamflow and the current year's streamflow, P is the mean annual precipitation, and Q is the mean annual streamflow. It has been found by Sankarasubramanian et al. (2001) that the streamflow elasticity is highly correlated with the ratio of standard deviations of the streamflow and precipitation when multiplied by the runoff ratio. This suggests that the streamflow elasticity is representative of the relative variability of streamflow and precipitation. In other words, streamflow elasticity can be considered a descriptor of inter-annual variability in precipitation and streamflow.

Snow Day Ratio

The snow day ratio (R_{SD} [-]) defined by the number of days that experience precipitation when the average daily temperature falls below $2^{O}C$ divided by the total days per year experiencing precipitation (Clark et al., 2006; Clark and Vrugt, 2006). This value represents a measure of seasonality in defining the degree to which spring snowmelt occurs. Since this value is also tied to the temperature and duration of a catchment below a temperature threshold, value also may affect biological processes.

$$R_{SD} = \frac{N_S}{N_P} \quad \text{Eq 7.}$$

Where N_S is the number of days of precipitation where the average daily temperature falls below $2^{O}C$ and N_P is the number of days of precipitation. A high value of snow day ratio means that a large amount of storage is provided as snow, and can be interpreted as a high degree of spring snow melt, and coincidentally, a large amount of intra-annual variation on how the catchment responds to precipitation events. A low value of snow day ratio means there is little to

no storage in the form of snow, and can be interpreted as a more regular catchment response throughout the year.

Recession Coefficient

The Recession Coefficient (b [-]) is the slope of the scatter plot created from plotting the mean streamflow vs. the slope of the recession curve in log-log scale (Brutsaert, 1977; Vogel and Kroll, 1992). The value of b carries with it information about the shape of the groundwater recession time series curve. In general, when looking at the hydrograph plotted at normal scale, a low value of b means there is a low curvature to the recession curve and a high value of b means there is a high curvature to the curve.

A value of 1 would dictate linear reservoir behavior, where the change in streamflow is directly proportional to the amount of streamflow. (Wittenberg and Sivapalan, 1999). A value of 2 is considered to share attributes with TOPMODEL aquifer behavior, with a much sharper decline within the recession portions within the hydrograph. This value carries with it information about the slope of the bedrock surface, which is usually assumed to be similar in slope to the ground surface. These values are calculated under the following conditions in this study as suggested in part by Allen et al. (1995)

- The streamflow is comprised solely on baseflow as calculated by using the Digital Filter Method
- At least 10 years of data is ideal in order to capture a enough data to calculate robust coefficients.
- Scatter plots exhibit a linear coefficient of fit value greater than 0.55 to ensure robustness in calculation. Catchments showing a value of less than 0.55 are not included in this study.

All streamflow recession curve data is separated and the value of b represented in the equations as beta) is calculated using the following equations. Technical details can be found in Brutsaert (1995).

$$\frac{(dQ)}{\Delta t} = f(Q_{average}) \qquad \text{Eq 8.}$$

Brutsaert concluded that a relationship exists between the change in daily streamflow, *dQ* and the two day mean streamflow, $Q_{average}$ as represented in Equation 8. When plotted against each other, a relationship normally exists.

$$\frac{(Q_2-Q_1)}{\Delta t} = \alpha \left(\frac{Q_2+Q_1}{2}\right)^b \quad \text{Eq 9.}$$

The equation representing this relationship is Equation 9 showing coefficients α and b representing flexibility in the slope and shape of the curve of fit to the data. This equation assumes that the curve of fit to the data will pass through the point (0,0), which is a reasonable assumption given that when the average streamflow measured is zero, the change in streamflow is also zero.

$$\log\left(\frac{dQ}{\Delta t}\right) = b \log\left(\frac{Q_2+Q_1}{2}\right) + \log(\alpha) \quad \text{Eq 10.}$$

When the data is log transformed, the log of both sides of Equation 9 transforms into Equation 10. The coefficient b becomes slope value and coefficient α becomes the intercept value. From Equation 10, a line of best fit can be calculated and the coefficients can be quantified.

$$b = \frac{\log\left(\frac{dQ}{\Delta t}\right) - \log(\alpha)}{\log\left(\frac{Q_2+Q_1}{2}\right)} \quad \text{Eq 11.}$$

In this study, the b coefficient is considered the recession coefficient (b). Equation 11 defines what the coefficient b equates to in relation to Equation 10.

An alternate approach to the one suggested in this paper is creating a master recession curve, using an exponential function to fit data along a composite time series. (Allen et al., 1995). Although it has been proven as a reliable way to calculate recession coefficients in the past, the process is time intensive and partly subjective, which makes it inappropriate for an empirical study of this scale.

3.2 Classification and Cluster Analysis

Hierarchical cluster analysis is an approach to group objects based on the similarity of their characteristic values. (O'Sullivan and Unwin, 2003). Euclidean distance is used to quantify distance between values or groups of values. The first step in the process is to calculate the mean and standard deviation of each data series The next step would be to create dimensionless values (standard scores or z-scores) using mean and standard deviation to each data point within the data series. Although the use of a z-score seems to suggest a normal distribution is needed, it is acceptable to use this method of non-dimensionalizing data series without a normal distribution. (O'Sullivan and Unwin, 2003). Each data point in a series represents a value relating to an object. In this case, a catchment is considered the object.

$$z_i = \frac{a_i - \bar{a}}{\sigma} \qquad \text{Eq 12.}$$

Where z_i is the z-score value at data point location i, a_i is the data value at data point location i, $\bar{a}$ is the mean of the data series, and σ is the standard deviation of the data series. Z-scores are calculated in order to non-dimensionalize and normalize all of the different data series into a similar range Ward's clustering method of calculating distances was chosen as the most appropriate method:

$$ESS = \sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2 \qquad \text{Eq 13.}$$

Where *ESS* is the error sum of squares, x_i is the score of the ith catchment. Catchments are grouped to form the smallest possible increase in variance at each step, and then each new group is treated as a single entity. Repeat the process until all objects are clustered together. To specify different clusters, a threshold is chosen depending on either a ratios of distances between objects or through process understanding of how the clusters are created.

Ward's method of clustering is the most useful type of combining values in order to preserve the amount of information within each point of data. (Ward, 1963). Other types of clustering such as single or complete clustering can lead to linking in which one group is created

in the first iteration and then keeps becoming larger and larger until all points are contained within the starting group. This makes identifying different clusters very difficult in that no distinct or differentiable groups form. In this analysis, a distance threshold of 5 in dimensionless space was used to group individual signatures, and a distance threshold of 18, 16, 13, and 7 were used to group the combination of signatures.

A dendrogram is a visualization tool used in conjunction with hierarchical clustering analysis to display data points and the value they represent. In this application, the catchments are considered points of data and the values they represent are values of each signature.

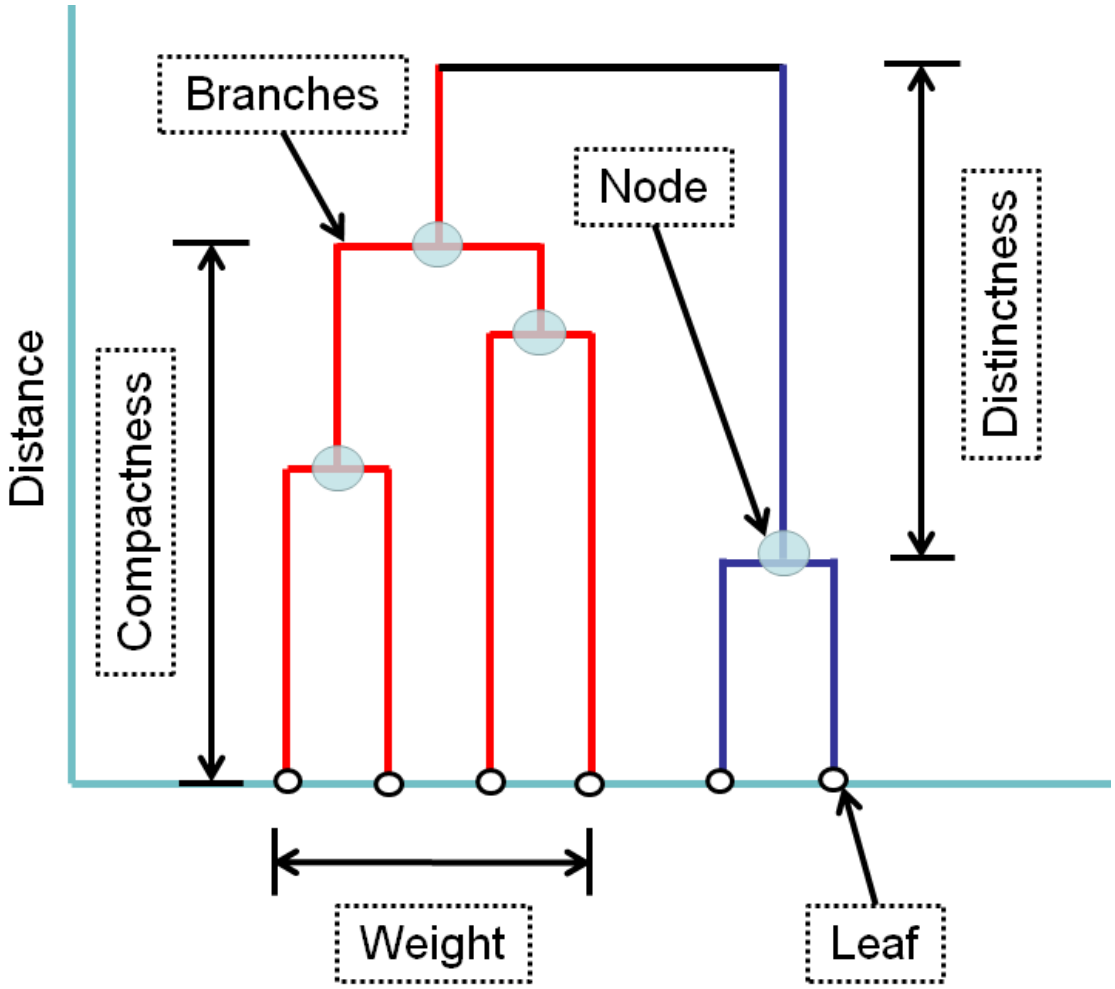


Figure 3-3. Schematic dendrogram introducing the descriptive terminology. The vertical axis represents a distance metric used in determining how close or far one catchment is to another catchment in signature space. The horizontal axis represents an order of catchments based on the

cluster analysis. The distance threshold is shown as a black horizontal line, whereas the groups are shown in red and blue.

An example of a dendrogram can be found in Figure 3-3. The vertical axis contains dimensionless Euclidean distance representing how far or close together points of data are in Attributes of a dendrogram can be described using the following terminology: Leaf, Branches, Node, Compactness, Distinctness, and Weight. A leaf represents the catchment of a signature value or collection of signature values. A branch describes the way that the different leaves connect to form groups. A node is the point at which the braches meet and a group is formed. The node is located at a certain distance value on the vertical axis representing how similar or dissimilar the points of data are. The threshold value can be considered the distance value in which nodes under this value form different groupings, or clusters. Compactness is a measure of distance represented by the position of the node. This defines how similar the data points within the group are according to their value. Distinctness is a measure of the distance between the present group node and the next node. A high value suggests a large dissimilarity between threshold values in which the current grouping and what would be the next grouping would exist. Weight is the number of data points or leaves contained within a group. A group with more data points than any other would be considered to have more weight.

Groups, or clusters, are formed by first setting a distance threshold (setting a distance value on the vertical axis). Any vertical portion of the dendrogram tree that crosses distance threshold is considered a separate group. Every group can then be measured by degree of compactness or distinctness by the location of its node. The number of groups in any cluster analysis is dependent upon the value of the selected distance threshold. Because this distance threshold changes for every application, several different distance thresholds can be applied and analyzed. From the resulting clusters and the signature values that control them, a classification level can be chosen.

An alternative approach to hierarchical clustering is k-means nonhierarchical clustering which has also been proven to be a semi-reliable way of forming clusters (Lagacherie et al., 1997; O'Sullivan and Unwin, 2003). However, this approach is difficult to visualize in six-dimensional space needed in this study and it does not display the relationship and progression of clusters from one to another as the number of clusters decrease, or as the threshold value increases.

Chapter 4

Catchments and Data

4.1 Catchment Characteristics

A total of 280 US catchments were used in this study spanning the eastern half of the United States as can be seen in Figure 4-1. The catchments used in this study exist in a similar range (east of the Rocky Mountains) to a study performed by Milly (1994).

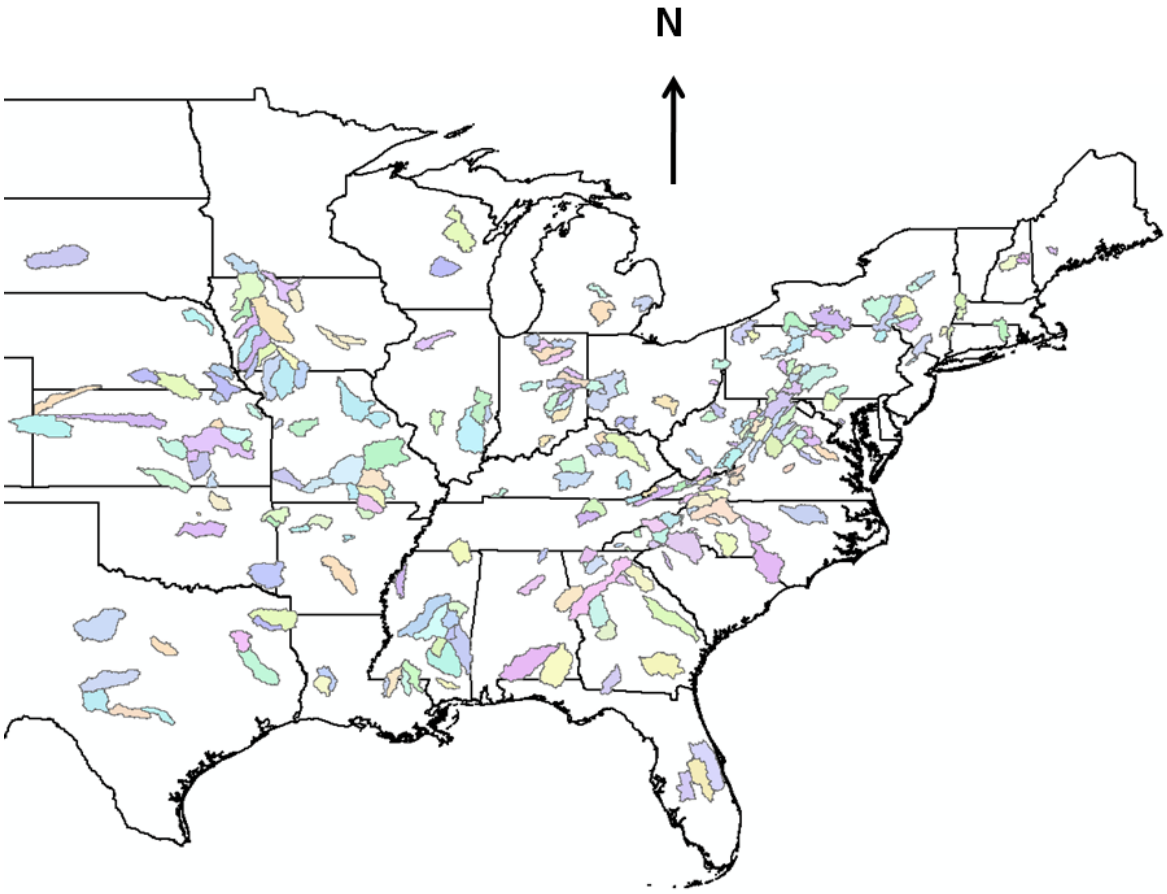


Figure 4-1. Map showing the 280 US catchments delineated through use of ArcGIS Spatial Analyst Toolbox and topography data. Catchments were selected from the MOPEX database and used for this study

The data used in the analysis was made available by the MOPEX project (Duan et al., 2006). MOPEX provides a collection of time series and static catchment information derived from other sources. Topographic data was supplied by the USGS in the form of 30 meter resolution digital elevation models (DEM) and DEM processing was performed using the ArcGIS Spatial Analyst toolbox. Characteristics Information provided by this dataset includes area, vegetation type (UMD system), soil type at 150 cm and 250 cm, greenness fraction, NDVI 5 week maximum gradients. Catchment boundaries from processed DEM data were used to calculate catchment area and boundary lengths. Catchments range in size from 175km^2 to 9000km^2, lengths range from 404km to 14.3km. The catchments cover type 1 eco-regions 5, 8 and 9, which are defined as Northern Forests, Eastern Temperate Forests, and Great Plains,

respectively. A summary of the catchments, names, and some attributes used in this study can be found in Table A-1 and A-2.

4.2 Time-series

Daily potential evapotranspiration contained within the MOPEX dataset is calculated based on NOAA Freewater Evaporation Atlas. This approach finds the long term values of evaporation and fit the averages to a unit curve specified by NOAA to capture intra-annual distributions of evaporation. This approach only estimates the seasonal energy forcing based on a generic curve and therefore should not be considered real data at the daily time scale. However, using this data to come up with mean annual evapotranspiration values is considered real data and is acceptable.

Precipitation information is provided by National Climate Data Center (NCDC) stations. Daily precipitation values were used in this study, however it is worth noting that all catchments in this study also show hourly precipitation data as well. Daily precipitation stations outnumber hourly precipitation stations by a factor of 4 to 1. In order to attain hourly data for all catchments, daily precipitation information is disaggregated to hourly data using the nearest hourly recording station. Minimum accepted precipitation gauge density found within each catchment follows the equation

$$N = 0.6A^{0.3} \qquad \text{Eq 14.}$$

Where N is the number of precipitation gauges and A is the area of the catchment (km^2) (Duan et al., 2006). A certain minimum number of precipitation gauges according to equation 14 are required in order to reasonably capture heterogeneity of storm events in order to come up with reliable spatially averaged values for a catchment. 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.

Daily streamflow data originates from USGS gauges. Only catchments that represent a minimum required streamflow data time series of a length of 40 years were considered to ensure sufficient data would be provided. In this study, 10 years of data was used between 1949 and 1958. The remaining 30 years of data will be used in future studies which are discussed in more detail in Chapter 6.

Chapter 5

Results

5.1 Individual Signature Explanation

Low values of **runoff ratio** generally indicate a catchment with a high actual evapotranspiration and low streamflow and vice versa. Milly (1994) found that in catchments east of the Rocky Mountains, seven dimensionless numbers are predictors of runoff ratio. These dimensionless numbers can be described as the dryness index (PE/P), plant available water in the soil divided by precipitation amounts (Ψ/P), and five (5) different measures of seasonality and variability in climate and soil. This suggests that both climate and soil properties may control the value of R_{QP}. Catchments that experience higher PE/P and Ψ/P ratios coincide with catchments with high evaporation and therefore low runoff ratios.

Figure 5-1(a) shows a catchment in Missouri that has a PE/P ratio of 1.29, while Fig. 5-1(b) shows a catchment in Indiana that has a PE/P ratio of 0.829. The differences in available energy and moisture are reflected in their runoff ratios.

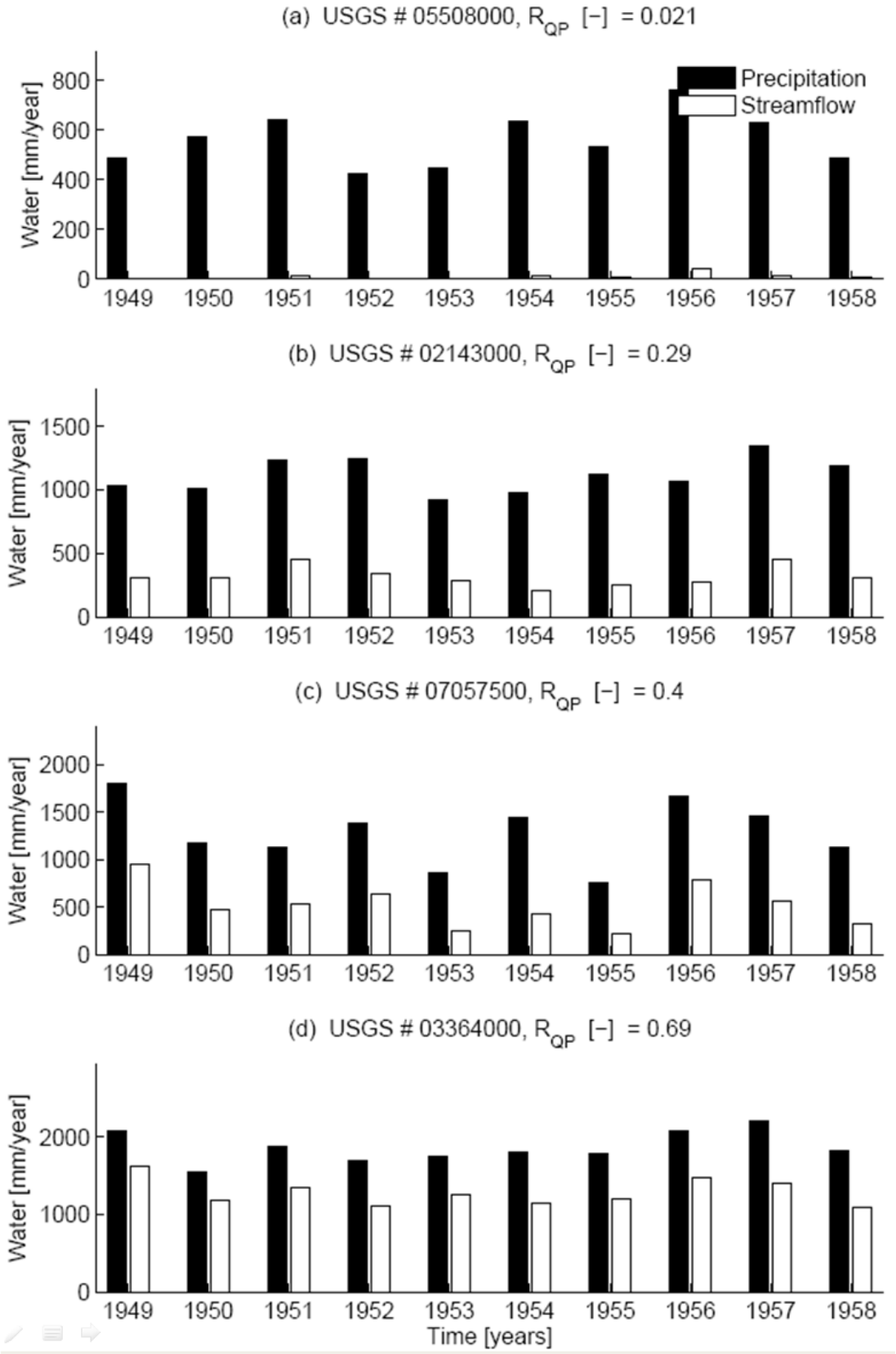

(a) USGS # 05508000, R_{QP} [−] = 0.021
Precipitation
Streamflow
Water [mm/year]
0
200
400
600
800
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
(b) USGS # 02143000, R_{QP} [−] = 0.29
Water [mm/year]
0
500
1000
1500
(c) USGS # 07057500, R_{QP} [−] = 0.4
Water [mm/year]
0
500
1000
1500
2000
(d) USGS # 03364000, R_{QP} [−] = 0.69
Water [mm/year]
0
1000
2000
Time [years]

Figure 5-1. Bar graph of annual precipitation and streamflow of four selected catchments as temporal graph. Plots are sorted from (a) lowest to (d) highest runoff ratio (streamflow/ precipitation) available in the dataset. Graph (a) shows the catchment with the lowest value for R_{QP} (0.02). Graph (b) shows the catchment that has the 33rd percentile value of R_{QP} (0.29). Graph (c) shows the catchment that has the 66th percentile value of R_{QP} (0.4). Graph (d) shows the catchment that has the highest value of R_{QP} (0.69).

Slope of the FDC with high values convey the variability of streamflow values over the linear portion of the curve plotted in semi-log space between the 33rd and 66th percentile streamflow exceeded. It is worth noting that the catchment connected to USGS gauge 06897500 located on the Grand River near Gallatin, MO fails to exhibit all potential variability because of no flow days contributing to over 20% of the time.

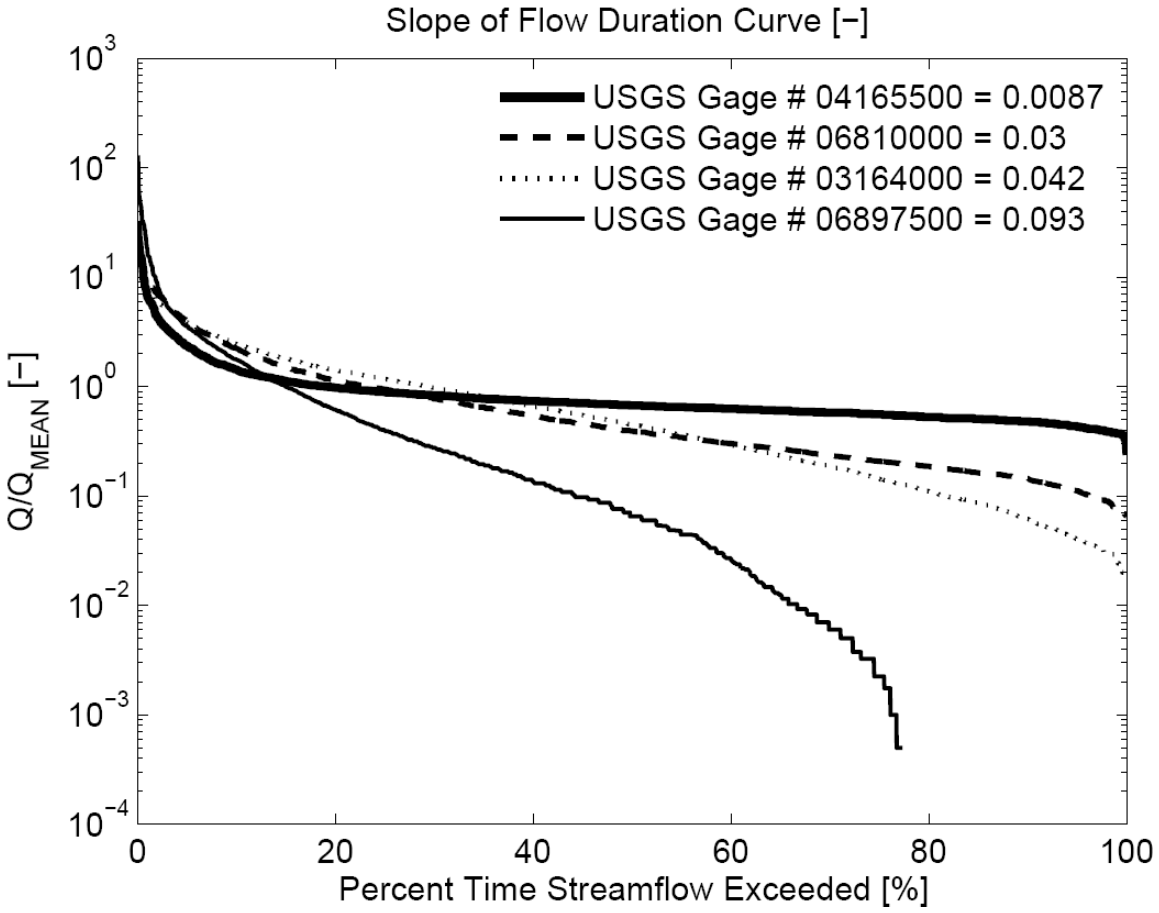


Figure 5-2. Semi-log plot of flow duration curves of four selected catchments. The vertical axis represents streamflow values normalized by the mean streamflow value of each catchment. FDCs of catchments are sorted from (a) lowest to (d) highest S_{FDC} available in the dataset. Gage # 04165500 (—) shows the catchment with the lowest value for S_{FDC} (0.0087). Gage # 06810000 (- - -) shows the catchment that has the 33rd percentile value of S_{FDC} (0.03). Gage # 03164000 (···) shows the catchment that has the 66th percentile value of S_{FDC} (0.042). Gage # 06897500 (—) shows the catchment that has the highest value of S_{FDC} (0.093).

FDCs shown represent highest and lowest values in dataset, as well as two intermediate values. S_{FDC} is controlled by how direct the flow path is to the catchment outlet. Physical properties that control the value of SFDC would affect the daily variability of flow and the way water is filtered through the catchment. This may depend on such properties like slope/topography/channel network of the catchment and soil storage capacity. A flow path that is less interrupted shows a greater variability in flow and therefore a steeper slope of the flow duration curve and a "flashy" response. The flow path with more interruptions shows less variability in flow and therefore a shallow flow duration curve and a "filtered" response.

Low values of **streamflow elasticity** indicate a catchment that is relatively insensitive to changes in precipitation input between years. In other words, a change in precipitation between years results in a smaller change in streamflow. High values indicate a catchment that is relatively sensitive to changes in precipitation input between years In other words, a change in precipitation between years results in a larger change in streamflow. A S_{FE} of 1 indicates a linear relationship between precipitation and streamflow. Values of S_{FE} were calculated using the median S_{FE} value.

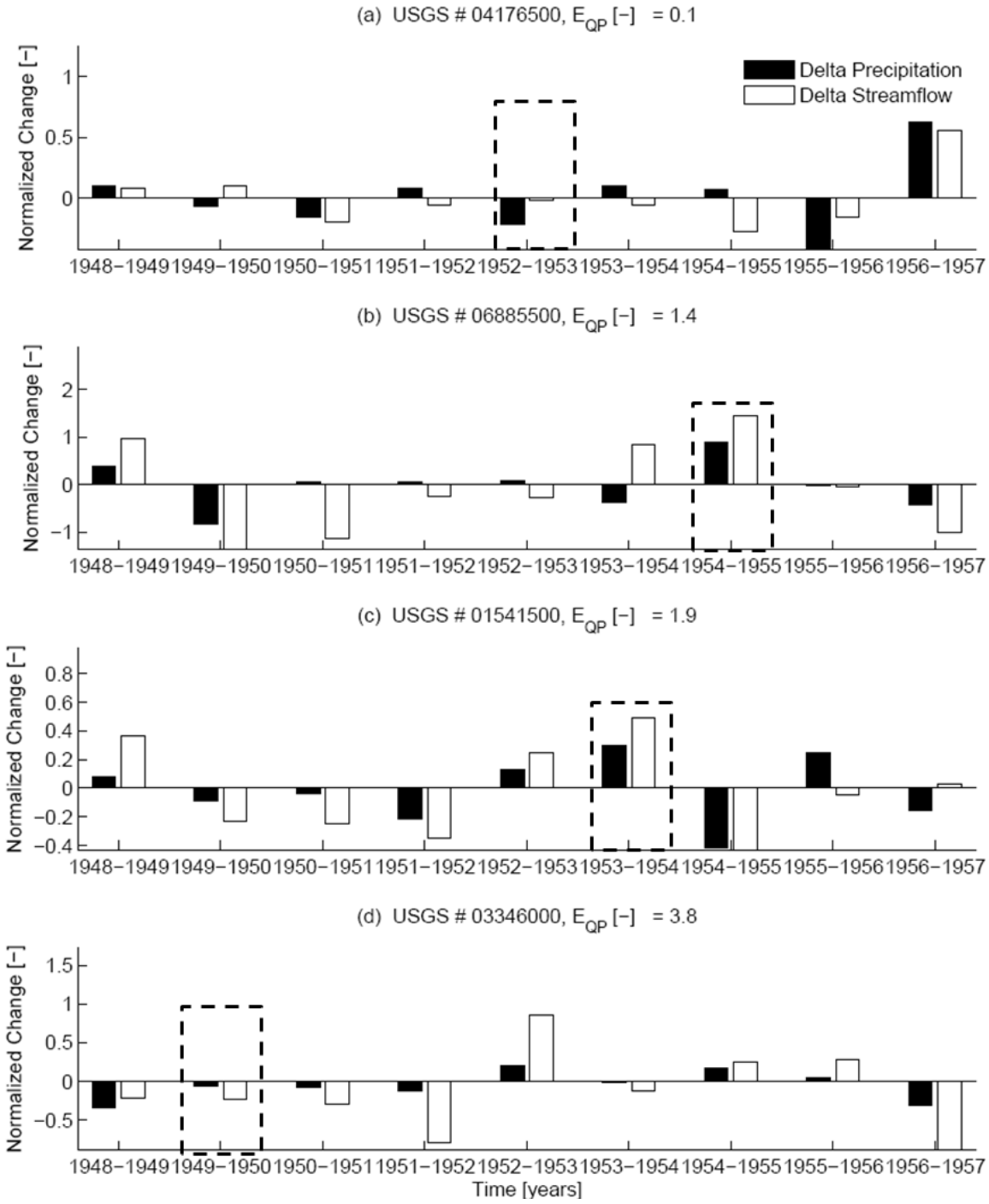


Figure 5-3. Bar graphs of differences in inter-annual precipitation and difference in inter-annual streamflow of four selected catchments normalized by mean precipitation and mean streamflow,

respectively. The median value for each gauge is highlighted by dotted black boxes. The ratio of these variables represents streamflow elasticity. Catchments are sorted from (a) lowest to (d) highest values of streamflow elasticity in the dataset. Graph (a) shows the catchment with the lowest value for E_{QP} (0.1). Graph (b) shows the catchment that has the 33rd percentile value of E_{QP} (1.4). Graph (c) shows the catchment that has the 66th percentile value of E_{QP} (1.9). Graph (d) shows the catchment that has the highest value of E_{QP} (3.8).

Figure 5-3(a) shows a catchment (04176500) located in Michigan that has a water demand that exceed available water during the median years both from vegetation, groundwater, and storage capacity as well as energy from climatic forcing show insensitive streamflow response to precipitation. Figure 5-3(b) shows a catchment (03346000) located in Illinois that has water demands that are being met and exceeded by the quantity of water during the median years show highly sensitive streamflow response to precipitation.

Low values of **base flow index** indicate a high amount of quick flow exiting the catchment through fast flow paths with a small amount of baseflow contributions between storm events as shown by a large portion of streamflow volume contributing during or just after a storm event. High values indicate a high amount of groundwater contribution to the streamflow and therefore slow flow paths with high contributions between storm events as show by a large portion of streamflow volume contributing between storm events. When exploring signatures to use, α as defined by Equation 10 is generally highly correlated with I_{BF}. According to studies performed by Vogel and Kroll (1992) and Kroll et al. (2004) using 23 catchments in Massachusetts, hydraulic conductivity (K_{sat}), porosity, drainage density, and groundwater slope were found to be possible descriptors of this alpha coefficient. According to a study performed by Beighley et al. (2005) studying 4 catchments in southern California, this alpha coefficient was found to be controlled primarily by hydraulic conductivity. Based on evidence from these studies, soil properties may be possible controls of the I_{BF}.

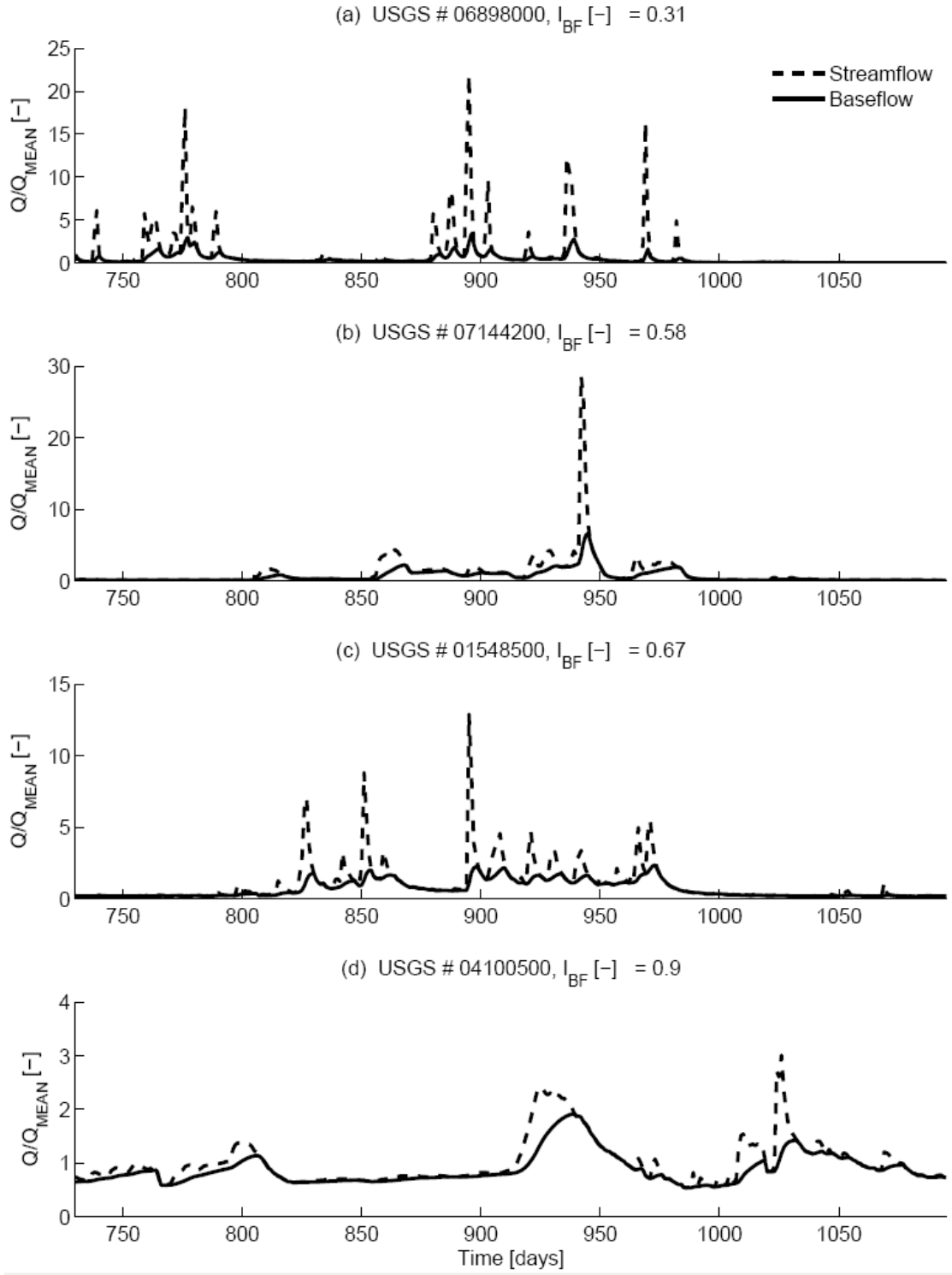


Figure 5-4. Plot of hydrographs and baseflow separation over the course of one (1) year for four (4) selected catchments categorized by base flow index. Catchments are sorted from (a) lowest to (d) highest values of I_{BF} for the dataset. Days are plotted from the third year of data using the

hydrologic year starting on October 1st (day 731) and ending on September 30th (day 1095). Graph (a) shows the catchment with the lowest value for I_{BF} (0.31). Graph (b) shows the catchment that has the 33rd percentile value of I_{BF} (0.58). Graph (c) shows the catchment that has the 66th percentile value of I_{BF} (0.67). Graph (d) shows the catchment that has the highest value of I_{BF} (0.9).

Figure 5-4(a) shows a catchment (06898000) located in both Iowa and Missouri represents a low base flow index. This catchment's dominant soils within the top 1.5 meters are found to be silty-clay-loam and clay, which are found to have low hydraulic conductivity. While Figure 5-4(d) shows a catchment (04100500) located in Indiana represents a high base flow index. This catchment's dominant soil within the top 1.5 meters is found to be loam, which has been found to have high hydraulic conductivity 3 to 4 times greater than soils found within the Iowa/Missouri catchment (Clapp and Hornberger, 1978).

Low values of **snow day ratio** reflect that little or no precipitation falls as snow or sleet. High values reflect seasonal snowfall and therefore periodic streamflow release during spring thaw exhibiting the storage of precipitation as snow during the winter months. When the mean air temperature rises above about 0^{o} C, a large stream response is seen. An example of this streamflow delay/ seasonality of flow can be seen in Figure 5-5. Snow is stored within the catchment until it is released through spring melt. The relative amount of snow stored within the catchment increases as the snow day ratio increases.

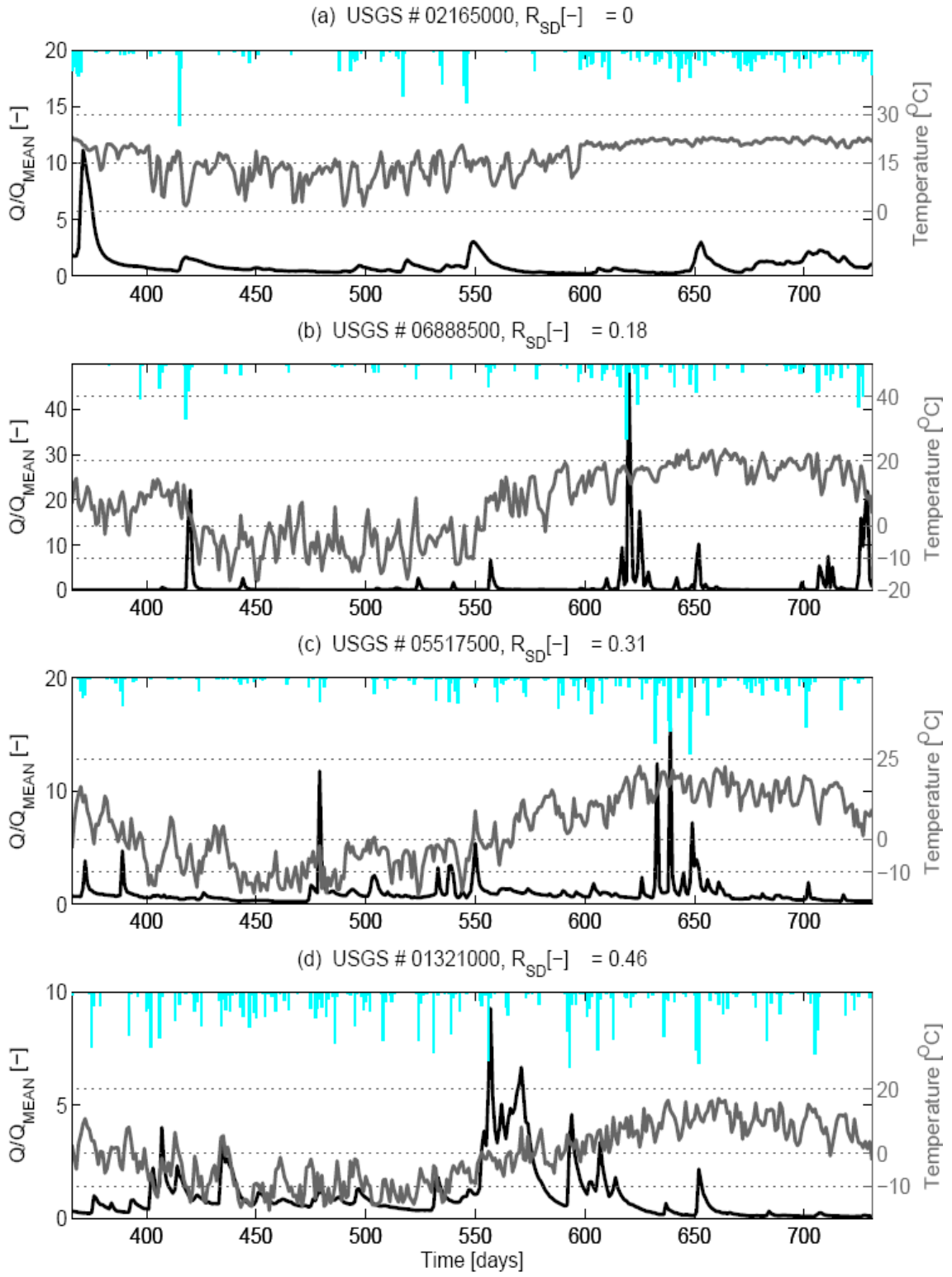


Figure 5-5. Hydrographs, average daily temperatures, and hyetographs of four selected plotted in one (1) graph per catchment. Catchments are sorted from (a) lowest to (d) highest values of R_{SD}.

Middle graphs show two in-between conditions. Days are plotted from the second year of data using the hydrologic year starting on October 1st (day 366) and ending on September 30th (day 730). Graph (a) shows the catchment with the lowest value for R_{SD} (0). Graph (b) shows the catchment that has the 33rd percentile value of R_{SD} (0.18). Graph (c) shows the catchment that has the 66th percentile value of R_{SD} (0.31). Graph (d) shows the catchment that has the highest value of R_{SD} (0.46).

Figure 5-5 presents catchments in ascending values of the snow day ratio. Catchment 02165000 shown in Figure 5-5 (a) is located in South Carolina and experiences no snow. This catchment experiences the lowest value of snow day ratio possible, as can be seen by temperature values that approach but never drop below 2^{o} C. A relatively regular streamflow response to precipitation events is displayed. Catchment 01321000 represented as (d) is located in New York and experiences snow 46% of the time. Streamflow response has little to no connection with precipitation events before day 550 when the temperature is well below 2^{o}C. Shortly after day 550 when the temperature rises above 0^{o}C, a very large streamflow response is seen until day 580 when streamflow response becomes more regular in relation to precipitation events.

Low values of the recession coefficient indicate a slow, linear release of groundwater into the stream as baseflow, indicated above as a small slope value. The line of best fit can be seen to have a relatively constant value of dQ to $Q_{mean.}$ High values indicate a fast, variable release of groundwater into the stream as baseflow, indicated above as a steep recession. The line of best fit can be seen to have a relatively steep slope.

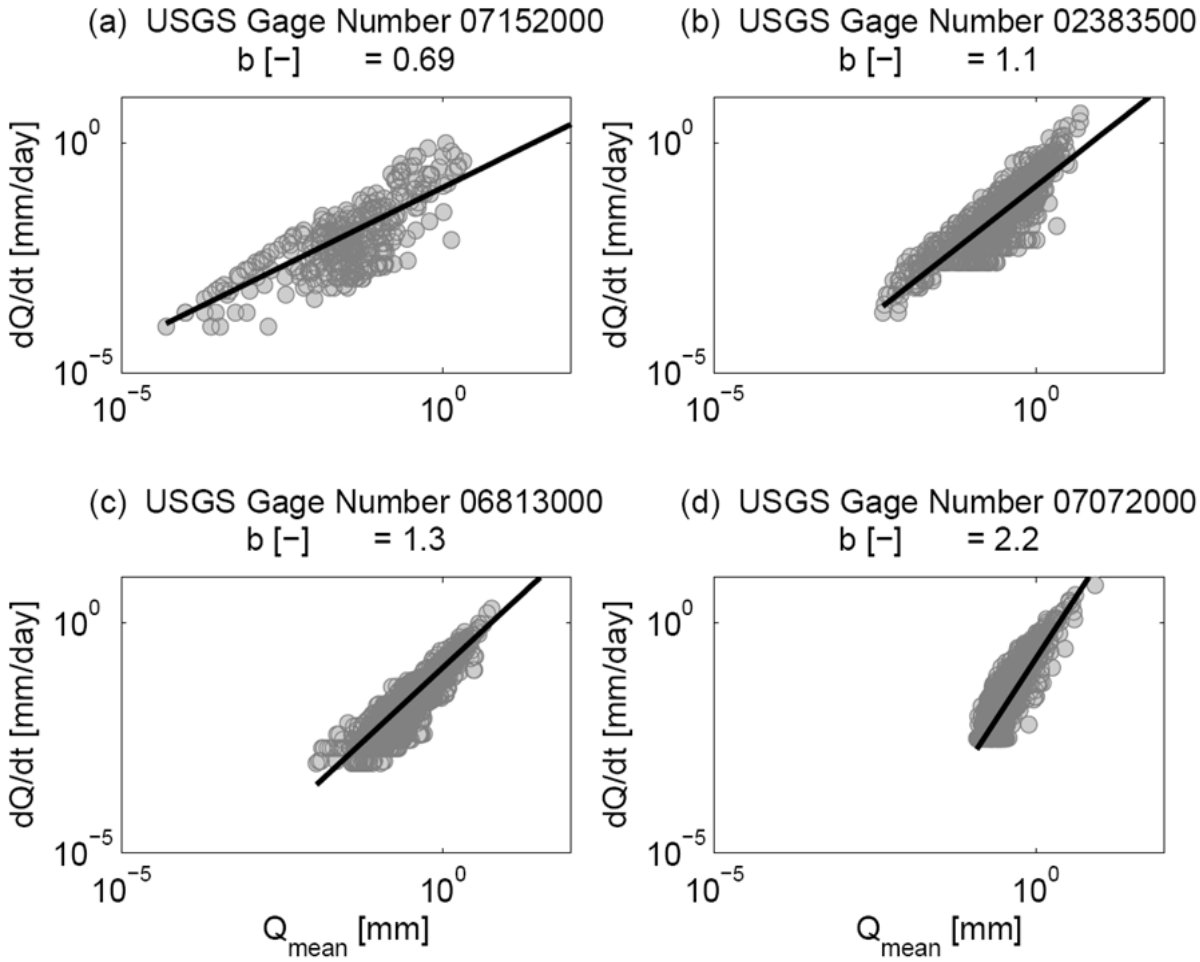


Figure 5-6. Scatter plot of averaged streamflow vs. the change in streamflow at the daily scale. Catchments are sorted from (a) lowest to (d) highest values of b, which is represented as the slope of the line of best fit in each plot. Graph (a) shows the catchment with the lowest value for b (0.69). Graph (b) shows the catchment that has the 33rd percentile value of b (1.1). Graph (c) shows the catchment that has the 66th percentile value of b (1.3). Graph (d) shows the catchment that has the highest value of b (2.2).

A study performed by Beighley et al. (2005) using 4 catchments located in southern California concluded that ground/bedrock slope is a physical predictor of the recession coefficient, b. This study was very small in scope, but offers a possible control of the value of the recession coefficient. Little information exists relating this signature to possible physical controls because most studies that use this coefficient assume the value of 1 which is defined as a linear

reservoir system (Vogel and Kroll, 1992). As seen in Figure 5-6, the linear reservoir assumption can be very inaccurate. Figure 5-6(a) shows a catchment (07152000) that is located in Kansas and Oklahoma represents a low value of the recession coefficient. Based on topography information from USGS DEM data, an average slope value of 0.19% was calculated. Figure 5-6(d) shows a catchment (07072000) located in Missouri and Arkansas with a high value of the recession coefficient. Topography information of this catchment shows an average slope value of 0.37%, which is twice the degree of slope experienced in the previous catchment.

5.2 Correlations and Spatial Patterns

Key signatures were chosen that were relatively uncorrelated (C_{lin} and C_{SR} below 0.45 with 1 exception) and represent a large distribution in values. I_{BF} and S_{FDC} showed the highest correlation value (about 0.66), however exhibit different correlation values when compared to the other four signatures, suggesting some difference in hydrologic meaning. I_{BF} is the measure of water contributing to the streamflow from the groundwater table. S_{FDC} is largely dependent on the path the water takes (longer/slower path results in low slope values). I_{BF} is a measure of one possible path that may be dominant in many catchments, however it is not the only possible path and therefore both signatures potentially explain different processes.

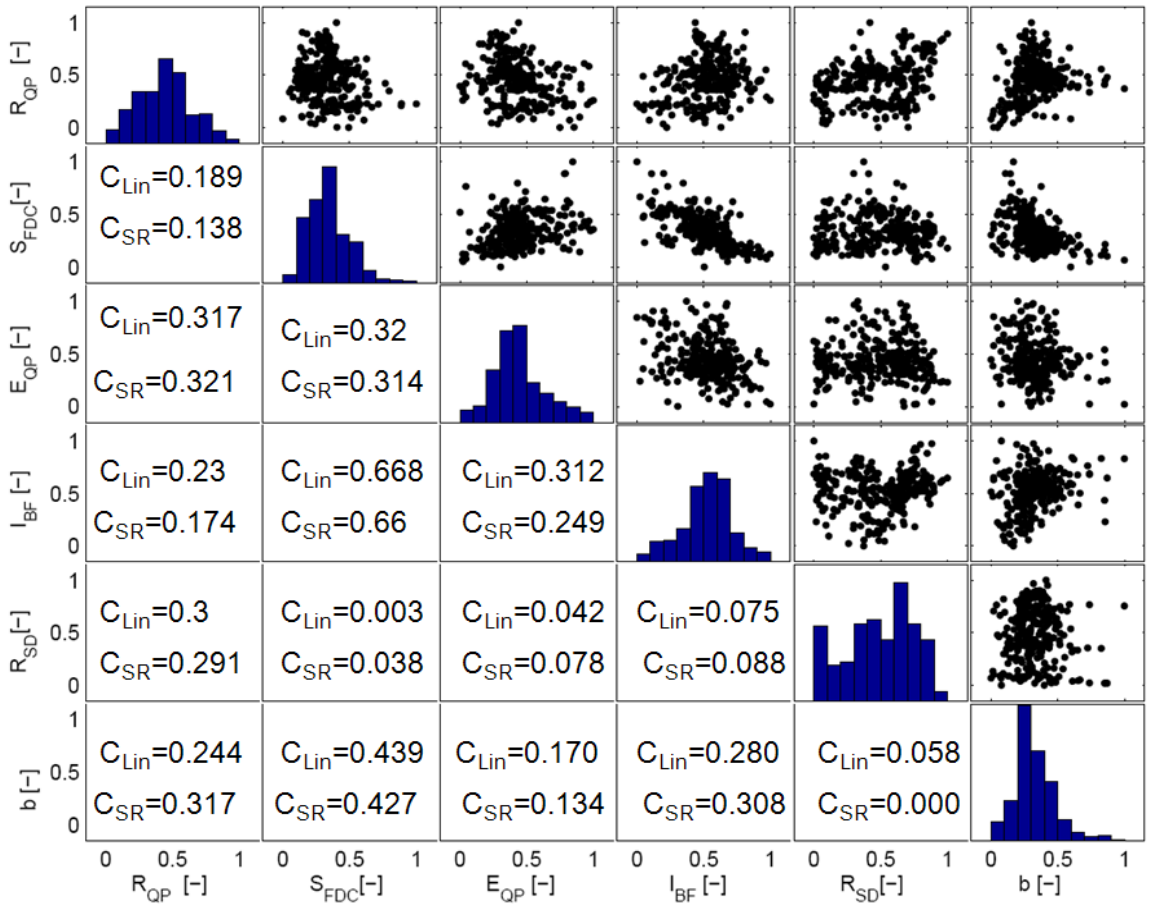


Figure 5-7. Multi-variable scatter plots and single-variable histograms of catchment signatures. Values shown indicate linear (C_{Lin}) and non-linear (C_{SR}) correlations. C_{SR} is the Spearman rank correlation coefficient.

As an example, the recession coefficient (b) was used in place of α due to a very high correlation between α and I_{BF}. In addition, key signatures were selected based on the amount of variability in signature ranges and distribution shapes. R_{QP}, S_{FDC}, and E_{QP} show a normal distribution with a light to moderate positive skew. The recession coefficient, b, shows a normal distribution with a strong positive skew. The baseflow index, I_{BF}, shows a slight negative skew in its relatively normal distribution. R_{SD} shows a varied distribution wish a slight negative skew.

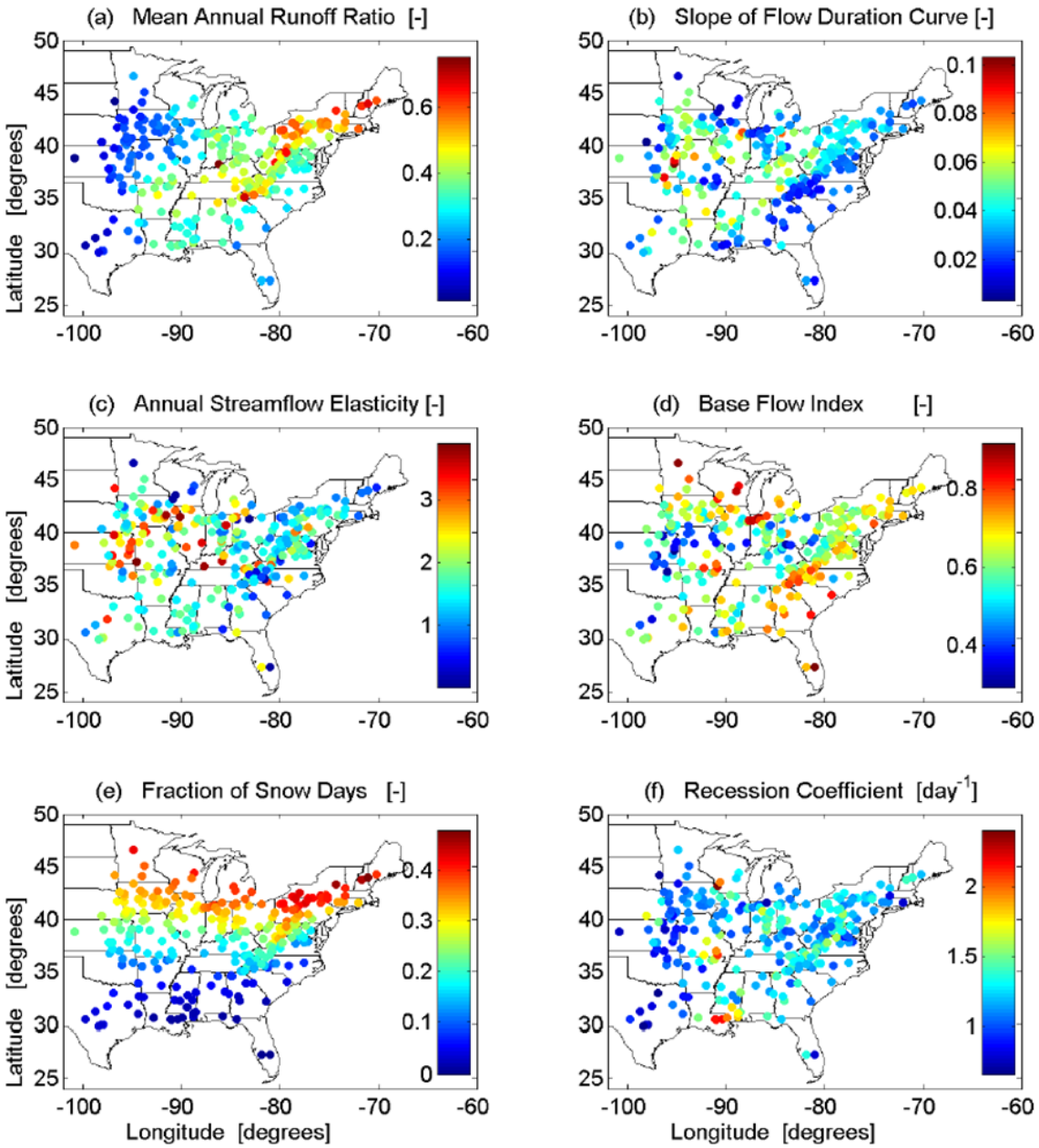


Figure 5-8. Each map shows the spatial distribution of catchment signatures at each stream gauge location, represented as a colored dot. The color of the dot corresponds to the colorbar next to each map. The color bar at each map shows the ranges of signatures used in this study. Plot (a) shows the spatial distribution of mean annual runoff ratio (R_{QP}). Plot (b) shows the spatial distribution of slope of the FDC (S_{FDC}). Plot (c) shows the spatial distribution of streamflow elasticity (E_{QP}). Plot (d) shows the spatial distribution of base flow index (I_{BF}). Plot (e) shows the

spatial distribution of ratio (or fraction) of snow days (R_{SD}). Plot (f) shows the spatial distribution of the recession coefficient (b).

From the spatial distribution of catchment outlets represented by their signatures, Figure 5-8 shows some signatures exhibit spatial patters while other signatures exhibit little or no spatial patterns.

The spatial plot shows a high value of **runoff ratio** along the Appalachian mountain range, suggesting that catchment slope and depth to bedrock as possible controls. The value of runoff ratio decreases as catchment distance from the Appalachian mountain range increases, and catchments seem to exhibit spatial similarity with respect to R_{QP}. This spatial distribution of runoff ratio values within catchments agrees with findings of Sankarasubramanian and Vogel, (2003). The spatial plot shows the lowest values of **slope of FDC** located on the southeastern side of the Appalachian mountain range. Slightly higher values are concentrated on the northwestern side of the Appalachian mountain range. The rest of the catchments show little to no spatial relationship with respect to S_{FDC}. Moderate to high values of **streamflow elasticity** are generally found in the Midwest section of the USA, however spatial proximity show weak relationships to similarity in E_{QP}. The majority of catchments with low **base flow index** are located around 40^{o} latitude. **Snow day ratio** shows strong relationship to latitude, which in turn has a strong effect on energy forcing. The catchments with the highest values of the **recession coefficient** are located along the Mississippi River.

5.3 Individual Cluster Analysis

The individual cluster analysis was completed to initially explore the signature separation and relationships. A value of 5 in dimensionless Euclidian space was used as a reasonable cutoff to create initial clusters for each signature based on visual inspection of the distinctness and compactness values. The dendrogram and recursive pattern plots ordered by each individual signature and labeled (a) through (f) can be seen in Appendix as Figure A1. This figure simply shows the individual signature separation that occurs depending on which signature is used for clustering and ordering the catchments. The weight of each grouping in each signature ordering

shows similarities to the distribution of values for that signature as shown in Figure 5-7. Figure A1 also shows possible correlations between signatures as first discovered in Figure 2-1.

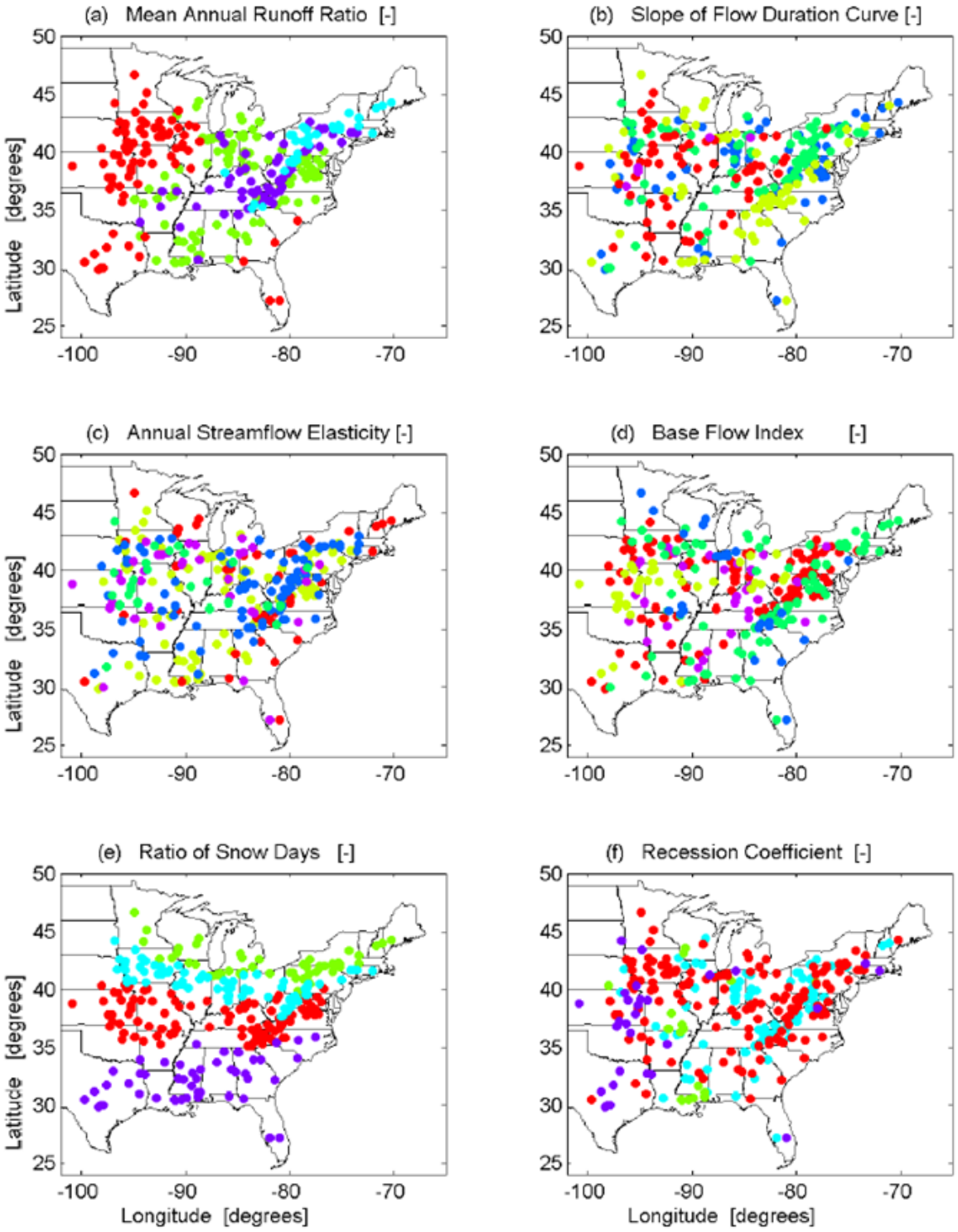
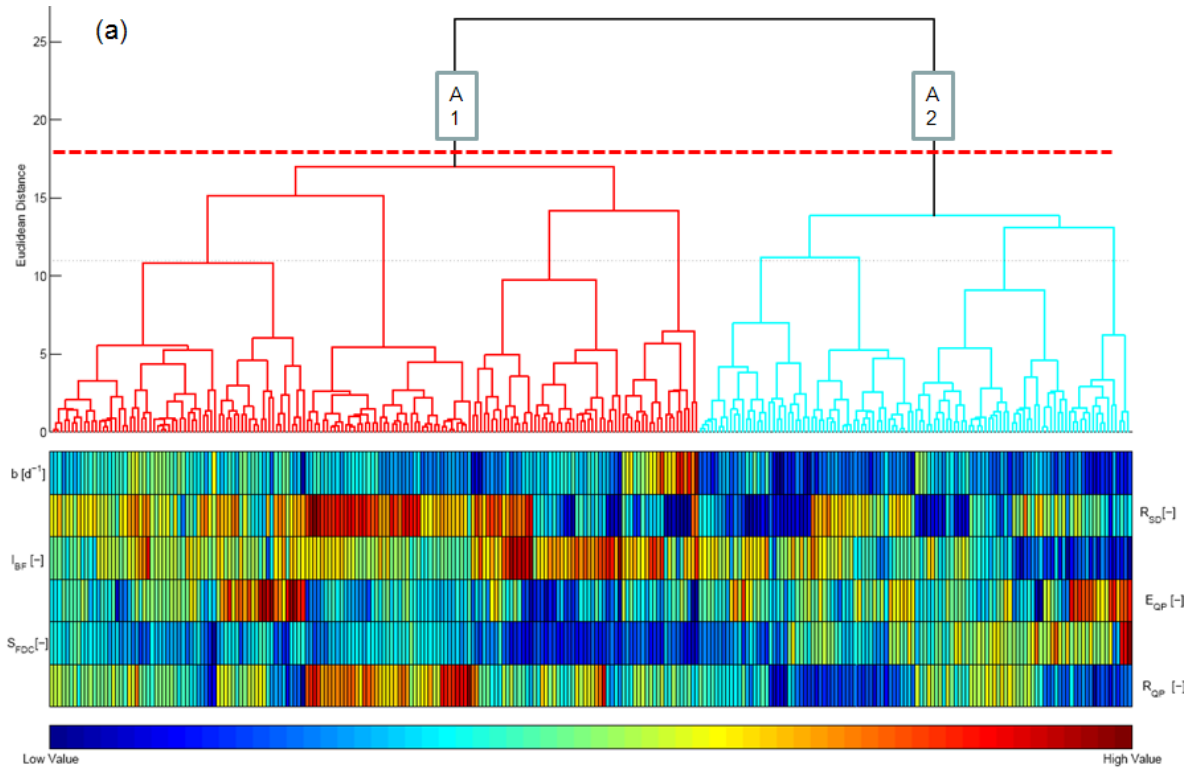


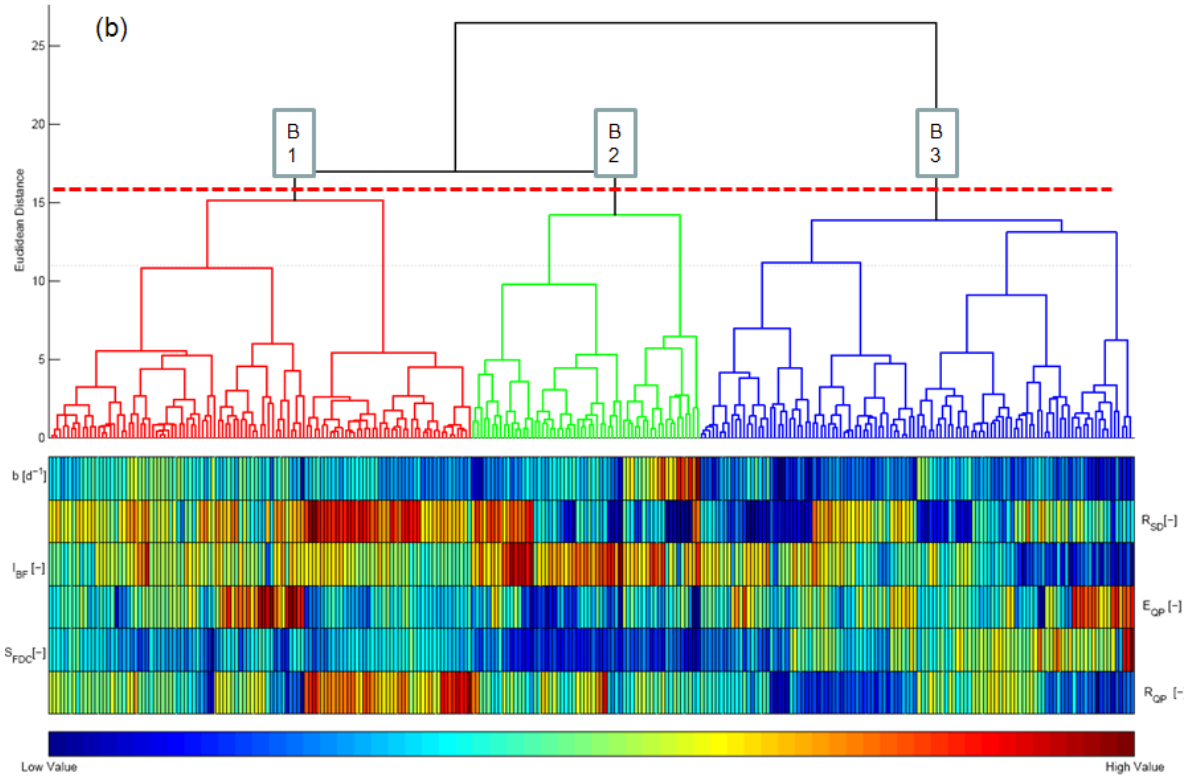
Figure 5-9. Spatial plot of groups ordered by each signature order from the individual cluster analysis. The colors represent the clusters that are created in Figure A1. Catchments ordered by the mean annual runoff ratio (a) results in 4 clusters. Catchments ordered by the slope of the FDC (b) results in 5 clusters. Catchments ordered by the streamflow elasticity (c) results in 5 clusters. Catchments ordered by the base flow index (d) results in 5 clusters. Catchments ordered by the ratio of snow days (e) results in 4 clusters. Catchments ordered by the recession coefficient (f) results in 4 clusters.

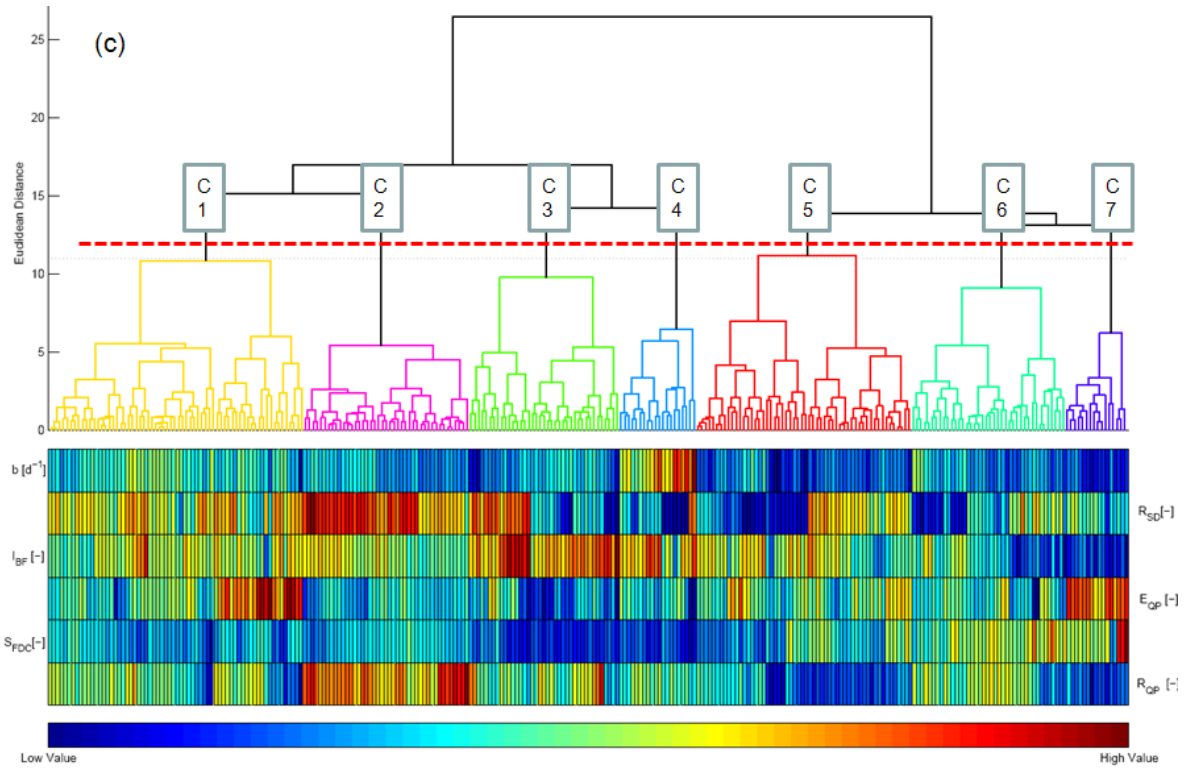
R_{QP} shows a strong relationship between catchments with high values and the Appalachian mount range. However, this area also experiences lower energy and higher precipitation than the area surrounding it most likely due to an orographic effect. Catchments with high values of S_{FDC} are located west of the Appalachian mountain range and are generally found away from large bodies of water. Streamflow elasticity shows no real spatial pattern. The base flow index shows a concentration of catchments with low values at latitude 40^{o}. The snow day ratio shows a strong relationship with latitude. Moderate values of b are found along the Appalachians, low values seem to be concentrated on the western most catchments, and the catchments with the highest value are located in close proximity to the Mississippi River.

5.4 Combined Cluster Analysis

A **combined cluster analysis** was performed combing all signatures into the cluster analysis to investigate levels of classification by chancing the distance threshold. The distance thresholds used were 18, 16, 12, and 7 each labeled as classification level A, B, C, and D, respectively. From these distance thresholds, 2 (A), 3 (B), 7 (C), and 11 (D) groups resulted (Figure 5-10). Controlling factors defined by the signatures were used to explain the nature of the catchments and of the groups to which they are assigned.


(a)
A 1
A 2
Euclidean Distance
25
20
15
10
5
0
b [d⁻¹]
I_BF [-]
S_FDC[-]
R_SD[-]
E_QP [-]
R_QP [-]
Low Value
High Value


(b)
B 1
B 2
B 3
Euclidean Distance
25
20
15
10
5
0
b [d⁻¹]
I_BF [-]
S_FDC[-]
R_SD[-]
E_QP [-]
R_QP [-
Low Value
High Value


(c)
Euclidean Distance
25
20
15
10
5
0
C 1
C 2
C 3
C 4
C 5
C 6
C 7
b [d⁻¹]
I_BF [-]
S_FDC[-]
R_SD[-]
E_QP [-]
R_QP [-]
Low Value
High Value

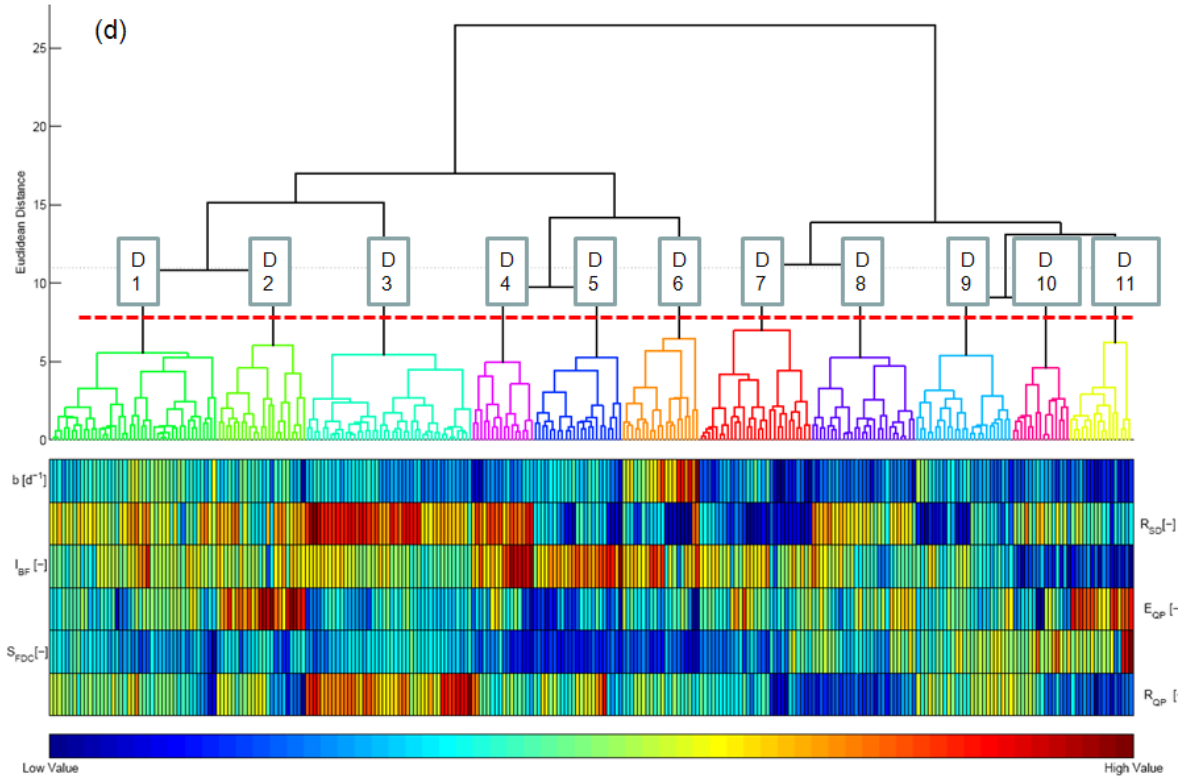

(d)
Euclidean Distance
25
20
15
10
5
0
D 1
D 2
D 3
D 4
D 5
D 6
D 7
D 8
D 9
D 10
D 11
b [d⁻¹]
I_BF [-]
S_FDC[-]
R_SD[-]
E_QP [-
R_QP [
Low Value
High Value

Figure 5-10. Recursive pattern plot ordered by the dendrogram of combined signature hierarchical cluster analysis. Classification level A (a) shows a distance threshold of 18 [-] and results in 2 groups. Classification level B (b) shows a distance threshold of 16 [-] and results in 2 groups. Classification level C (c) shows a distance threshold of 12 [-] and results in 2 groups. Classification level D (d) shows a distance threshold of 7 [-] and results in 2 groups.

Classification level A (CLA) results in the formation of two clusters. Group A1 shows the majority of catchments exhibiting high to moderate values of R_{QP}, E_{QP}, I_{BF}, R_{SD}, and b. Distance threshold of 18 was used to form groups. Group A2 contains catchments that exhibit high values of S_{FDC} and $E_{QP.}$ In both groups A1 and A2, catchment values exhibit a wide variation of catchment signatures resulting in groupings without very much explanatory power. A1 displays a higher degree of weight than A2, although A2 is more compact and displays a higher distinction than A1. Groupings with higher compactness are needed to explain the groupings.

Classification level B (CLB) results in the formation of three clusters. Group A1 from CLA is broken down into 2 groups labeled as B1 and B2. Grouping A2 carries over and is relabeled as B3. CLB still shows a relatively large amount of variation in signature values. A distance threshold of 16 was used to form groups. Group B1 shows catchments with high to moderate values of R_{SD}, along with moderate values of I_{BF}, and moderate to low values of b. R_{QP} and E_{QP} show a great deal of variability. This group consists of catchments controlled by R_{SD} to a large extent, and therefore can be attributed to latitude and air temperature. B1 seemed to be grouped using climatic descriptors. Group B2 shows high values of I_{BF}, low values of S_{FDC}, and relatively low values of E_{QP}. This group consists of catchments that are slow responding and results in streamflow that is primarily comprised of groundwater. Based on prior studies, it is possible that this group is controlled by physical characteristics. Group B3 contains moderate values for most signatures, although contained within this group are high values of S_{FDC} and E_{QP}, however these signature values are quite variable within this group. Group B3 shows the highest degree of compactness, controlled by its incorporation of mainly moderate values for most signatures. Groups B1 and B2 show very small distinction from one another as can be seen in the dendrogram. The vertical distance from where the groups are formed (at the corresponding node) and the formation of what would be the next group (Group A1) is very small compared to the compactness B1 and B2 exhibit. B3 also carries the most weight, however offered little explanatory power.

Classification level C (CLC) results in the formation of seven clusters. Group B1 and B2 were each split into two groupings and B3 was split into three groupings. Distance threshold of 12 was used to form groups. Group C1 contains catchments that represent high to moderate, varied values of E_{QP}, moderately low values of S_{FDC}, and moderate values of the other signatures. This group is represented by high to moderate E_{QP} and at the same time shows a filtered response, most likely due to a moderate amount of groundwater contribution of streamflow. A label for this group could be "Filtered/Sensitive." Group C2 contains catchments that represent and are controlled by high values of R_{QP} and R_{SD}. This grouping therefore is controlled by low energy/precipitation ratios and low temperature for a greater part of the year. This group can be labeled "Low Energy/High Seasonality." This group also exhibits the greatest distinction from the other groups. Group C3 contains catchments that represent high values of I_{BF}, low values of S_{FDC}, low values of b, and a large variation in values of R_{SD}. A large variation in R_{SD} signifies this group acting independent of temperature forcing. This group is controlled by I_{BF} and S_{FDC}. The filtered response represented by low values of S_{FDC} is possibly due to infiltration into the groundwater table. Catchments in this group exhibit low b values and suggest recession curves with a very small degree of shape, meaning a linear or slightly convex curve. This group can be labeled "Low Flow Regime Variability."

Group C4 contains catchments that represent high values of b, relatively high values of I_{BF}, and low values of S_{FDC}. This means that the recession curve is highly convex, most the streamflow is comprised of groundwater, and catchment response is flashy and unfiltered.. As in group 4, the controlling process for filtering the catchment response as represented through the S_{FDC} is though infiltration. This group can be labeled "High Flow Regime Variability." Group C5 contains catchments that represent moderate to moderately high values of I_{BF}, low values of R_{QP}, varied values of R_{SD}, moderate values of E_{QP}, low to moderate values of S_{FDC}, and low values of b. There exists a large variation in R_{SD} and therefore seasonality, but the ratio to which water exits the catchment as streamflow is relatively low. According to previous studies mentioned, it is possible that soil storage capacity at the surface is high, allowing time for evapotranspiration to take place. Because of this high soil storage capacity, I_{BF} is found to be moderate to moderately high. A label that can be used for this group is "Low Energy/Small Recession Curvature." Group C6 contains catchments that represent moderate RQP, moderately low b, moderate S_{FDC}. This group represents a moderate energy and a moderately flashy response. This group can be labeled "Moderate Flashy/Energy." Group C7 contains catchments that represent high values of E_{QP}, low values of I_{BF}, low values of b, low values of R_{QP}, and high values of S_{FDC}. This groups exhibits

mainly fast, flashy surface flow responses. High values of E_{QP} represent very low usage/storage potential while low values of R_{QP} a high energy/precipitation ratio. This group can be labeled "Flashy/Sensitive."

Classification level C is the first clustering level in which catchments can be properly described and labeled by the signatures that explain that group. The distance threshold used in this classification level can be considered the highest distance in which functional clusters are created. Catchments within those clusters start to show similar hydrologic signatures for use in describing each catchment hydrologically. This offers the greatest potential for flexibility and although more information can be derived from an additional classification level, the clusters that are created in classification levels with lower distance thresholds are less flexible and prone to being less robust.

Classification level D (CLD) results in the formation of eleven groups. Group C2, C4, and C7 all carry over to CLD due to their high degree of distinctness witnessed in CLC. Distance threshold of 12 was used to form groups. Groups D1 and D2 are a subset of C1. Like catchments contained within C1, catchments belonging to groups D1 and D2 experience moderately low values of S_{FDC}. These groups differ in that catchments from group D1 exhibit moderate values of E_{QP} where group D2 exhibits very high values of E_{QP}. Extending the label from group C1, group D1 can be considered "Filtered/Moderately Sensitive" and group D2 can be considered "Filtered/Highly Sensitive." Group D3 is group C2 carried over to CLD, and therefore its attributes and label are the same as group C2. Groups D4 and D5 are a subset of C3 and therefore both contain catchments which exhibit high values of I_{BF} and low values of b. Group D4, in addition to these attributes, also experiences a high R_{SD} whereas catchments within group D5 experience starkly lower values of R_{SD}. There is also a slight differentiation of E_{QP} between groups D4 and D5. D4 catchments show moderate sensitivity whereas D5 contains insensitive catchments. The seasonality and temperature difference shown between D4 and D5 could possibly relate to the difference in sensitivity, since the differences first appear at the same classification level given the attributes first exhibited in C3. Possible descriptors of D4 and D5 are "Low Flow Regime Variability/High Seasonality" and "Low Flow Regime Variability/High Seasonality," respectively. Group D6 is group C4 carried over to CLD, and therefore its attributes and characteristics are the same as group C4. Groups D7 and D8 are a subset of C5 and therefore both contain catchments exhibit an attribute of "Low Energy/Small Recession Curvature".

Group D7, in addition to this attribute, exhibits a low R_{SD} whereas group D8 exhibits a moderately high R_{SD}. D7 also exhibits a moderately low R_{QP} whereas D8 exhibits a very low

R_{QP}. This difference is worth noting because group D3 shows a strong interdependence of R_{SD} and R_{QP}, while D7 and D8 show independence of these signatures. Group D7 could be labeled "Low Energy/Small Recession Curvature/ Low Seasonality." Group D8 could be labeled "Low Energy/Small Recession Curvature/ High Seasonality." Groups D9 and D10 are a subset of C6 and therefore both contain catchments that show "Moderate Flashy/ Moderate Energy" attributes. Group D9 also exhibits low R_{SD} and moderate I_{BF}, whereas group D10 exhibits moderate R_{SD} and low I_{BF}. This distinction suggests that under similar (moderate) flashiness and energy/precipitation conditions, seasonality in streamflow response may control the amount of groundwater recharge under moderate to low conditions. This suggests that for catchments found in D10 spring thaw could saturate the top layer of soil causing a smaller contribution to groundwater than in experienced by catchments belonging to group D9, which show a greater baseflow contribution tied to a low degree of seasonality. D9 could be labeled as "Moderate Flashy/ Moderate Energy/ Low Seasonality". D10 could be labeled as "Moderate Flashy/ Moderate Energy/ Moderate Seasonality." Group D11 is group C7 carried over to CLD, and therefore its attributes and description are the same as group C7.

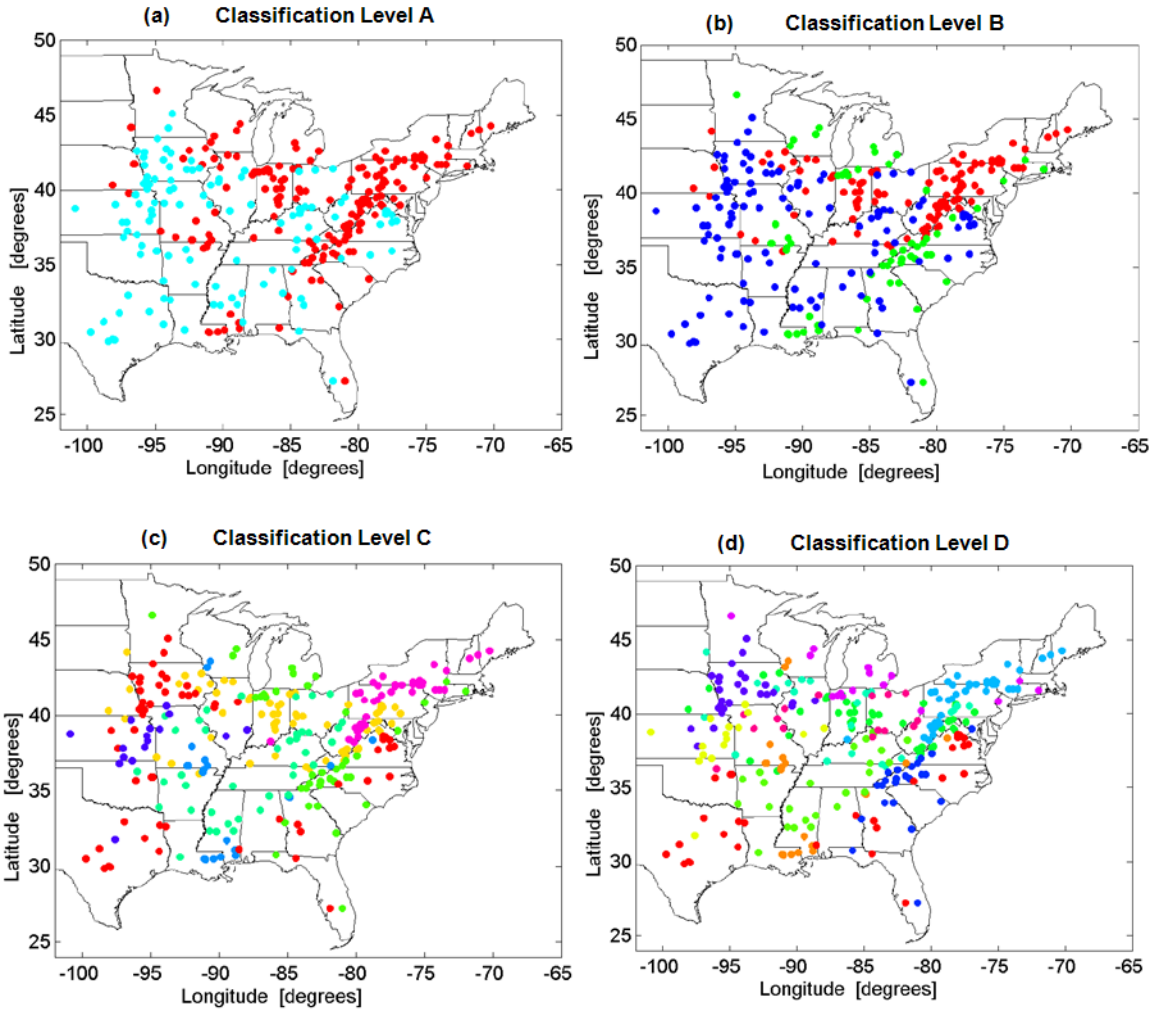


Figure 5-11. Spatial plot of groups created from each classification level. The color of each cluster represented in the plot matches the colors used in the dendrogram found in Figure 5-10. A spatial plot of catchments from CLA (a) results in 2 clusters, shown in red (A1) and light blue (A2). A spatial plot of catchments from CLB (b) 3 groups, shown in red (B1), green (B2), and blue (B3). A spatial plot of catchments from CLC (c) results in 7 groups. A spatial plot of catchments from CLD (d) shows a distance threshold of 7 [-] and results in 11 groups.

Figure 5-11(a) shows the spatial plot of the result of CLA. Group A1 consists of catchments that are largely located in the northeast and gradually mix spatially with catchments from group A2 as you move away from the northeast. Catchments that belong to group A2 are largely located in the west gradually mixing with catchments from A1 as you move eastward.

Figure 5-11 (b) shows the spatial plot of the result of CLB. Group B1 is located largely along the Appalachians and the northern part of the US. B1 was found to have largely climatic descriptors, which are defined by high values of R_{QP} and R_{SD}. Catchments belonging to group B2 are located across the US, although there is a concentration of B2 catchments located on the southern end of the Appalachians. B2 was found to be controlled mainly by physical descriptors. Catchments belonging to group B3 are located across the US with a concentration found in the western part of the catchments studied. There were no defining descriptors that were found to represented this group as a whole, but spatial relationships can be seen in the western catchments.

Figure 5-11 (c) shows the spatial plot of the result of CLC. Catchments from group C1 are located close to 40^{o} latitude, which is similar to what was found in the spatially plotted individual cluster analysis of I_{BF}. Group C2 is defined as having low energy and low seasonality. It is comprised of catchments that are located mainly in the northeast section of the US. Group C3 is defined as "Low Flow Regime Variability" and is found to be located on the eastern portion of the catchments studied. A large portion of these catchments are located at the southern end of the Appalachians. The low average slope that was used to describe these catchments may be related to the slope of the groundwater table at the connection to the stream network. Catchments found in this group can also be found in the northern most portions of the catchment study. Group C4 which is attributed with having high groundwater/bedrock level slopes show a clustering around the Mississippi River, defining steep groundwater gradients in proximity to the river. Group C5 which is defined as "High Soil Capacity" catchments cluster in the northwestern and southwestern parts (Midwest and Texas) of the catchment study, with only a few scattered catchments located in eastern portion of the catchment study. Group C6 which is defined as "Moderate Flashy/Moderate Energy" are clustered between what would be considered the Southeast and Midwest, spanning the area between them. Group C7 which is defined as "Flashy/Sensitive" is located in the most western portion of the catchment study, in what would be considered the beginning of the Great Plains.

Figure 5-11 (d) shows the spatial plot of the result of CLD. Groups D1 and D2 are found to be located within the same area as C1, however are dispersed within that area, suggesting little to no spatial relationships between D1 and D2. Groups D4 and D5 are located in the same locations as catchments in C3, however strong spatial relationships see to exist. Catchments part of D4 are found on the northern edge of the USA, whereas catchments of D5 are found along the southern portion of the Appalachians. Groups D7 and D8 are located in the same locations as catchments in C5, however catchments in D7 are located primarily in proximity to the coast line

and the Atlantic Ocean/Gulf of Mexico. Catchments belonging to D8 are located primarily in Iowa. Group D9 and D10 are located in the same locations as catchments in C6, however they show some degree of spatial clustering where catchments in D9 are located in the southern portion of catchments found in C6 and catchments in D9 are located in the western and northern portion of the catchments found in C6. This it most likely controlled by the difference in R_{SD} discussed earlier.

Chapter 6

Discussion

In total, over 40 potential signature values were calculated and explored. The signatures that were expected to show the highest information and lowest correlation were used to define aspects such as water balance (R_{QP}), inter-annual variability and sensitivity (daily time scale - S_{FDC}, yearly time scale-E_{QP}), intra-annual variability (R_{SD}), partitioning into surface and groundwater flow (I_{BF}), and release from groundwater storage (b). Controls for each signature were largely independent, however in the case of signatures I_{BF} and S_{DFC}, show that the majority of S_{FDC} values could be explained to some degree by values of I_{BF}. This was largely due to the fact that large contributions to and from the groundwater would filter the catchment response, resulting in a shallow slope of the FDC. However, variations and trends showing low I_{BF} and low S_{FDC} were also found and suggest other partitioning functions resulting in a filtered response. This may be due to storm intensity/duration or surface flow impedance such as leaf litter or channel network distribution.

Hierarchical cluster analysis in conjunction with recursive patter plots was repeated and 4 different distance thresholds resulting in 4 different classification groups were used in order to understand the amount of potential information available at each one. The first grouping (CLA) consists of two groups which display a high degree of variability in signature values within each cluster, although group A1 seems to contain the highest values of R_{QP}, I_{BF}, b, and R_{SD}. Due to the high variability in signature values, CLA provides little to no specific information as to what controls each group, however general trends do exist (Figure 5-11). Interestingly, spatial proximity at this level of classification can be used to describe where the different catchment clusters are located. Spatial proximity is most likely explained by general trends in signature values that contain climatic characteristics, such as R_{QP} and R_{SD}. CLB resulted in three groups, group B1 representing catchments of similar climatic characteristics where as group B2

represented catchments of similar physical characteristics. However, B3 provided little to no functional explanation for the groups in general. CLC showed a higher degree of interaction between signatures leading to greater explanation of catchment functions. CLC was also the classification grouping at the highest distance threshold that offered enough information to differentiate and label each grouping based on hydrologic processes and functions. Cluster analyses C has the greatest potential of representing a robust classification due to the large amount of weight associated with each. CLD contains 11 groups and was used to extract higher amounts of information that was lost in CLC. D1 through D11 all offered the highest amount of information about function interaction between catchments. Through CLD, general trends from the strong relationships found between data set I_{BF}/S_{FDC} and R_{QP}/R_{SD} in other cluster sets, exceptions organized in groups were found adding noise to the trend and offering dominant processes and interactions that may have been minor in other catchments. No further cluster analysis was performed at a lower distance threshold because catchments below a Euclidian distance metric of 7 form groups that start to have very similar attributes, but would be located in different groups diluting the information that is available. CLD was found to be the most informative cluster analysis performed, but CLC showed a greater potential for being robust.

In general, the approach offered in this paper is a robust and information dense way of analyzing catchment signatures, functions, and the relationship and classification between catchments of multiple types of hydrologic processes. Hierarchical cluster analysis result in different clusters at different distance thresholds based on the clustering method used (ie. single linkage, complete linkage, ect.). However, the Ward's method of calculating distance metrics and clusters was found to be a robust approach that resulted in well defined clusters. Displaying 7+ dimensional signature information (6 Individual and 1 Group) through use of a dendrogram plotted in parallel with a recursive pattern plot enabled the concentration of all data needed to classify catchments by their signatures and dominant processes. Displaying the clusters that result from the cluster analysis in a spatial plot further aided in the understanding of physical or climatic controls associated with those clusters. Spatial proximity shows some relationship to hydrologic similarity in determining clusters of catchments that experience similar signature values (Fig. 5-11).

The next step to this process would be to regionalize the key signatures to confirm the function and process interaction that have been suggested in this study. From this step, a classification of physical and climatic properties would be able to be completed based on hydrologic similarity. Studies that have been done in the past assume that the signatures used are

representative of the processes to be described. This study lays the foundation for process understanding to be derived from the hydrologic signatures, which is likely to be more robust in creating a catchment classification system built on the foundation of hydrologic similarity.

Chapter 7

Conclusions

The objectives of this study were to [1] identify key signatures derived from the streamflow response and other properties, and [2] to define and distinguish controlling hydrologic functions between catchments formalized in a classification system. Six key signatures that represented separate catchment functions were defined and analyzed including Runoff Ratio, Slope of the FDC, Base Flow Index, Streamflow Elasticity, Snow Day Ratio, and Recession Coefficient. A classification system was created based on these signatures at different threshold values. Analysis using hierarchical cluster analysis in conjunction with high dimensional recursive pattern plots was shown to be a very useful method in extracting and understanding the information between catchment signatures and catchment functions. Classification level C uses a dimensionless Euclidian distance measure of 12 and a Ward's clustering method resulted in seven groups. These groups each represented different controlling signatures that allowed for a well defined classification system. When the distance threshold was lowered and more clusters were formed, more information was able to be extracted for some groups, but offered little or no additional information of other well defined groups. A distance measure of anything less than 7 would lead to degradation in the organization of information. It is therefore suggested that a distance threshold between 12 and 7 and number of groups between 7 and 11 is an acceptable classification scheme for the key signatures that were selected. This method of classification shows a robust, information rich approach to classify catchments, satisfying the original goals of this study.

The next steps in extending this study would be to [1] relate the physically based characteristics of the catchments to the key signatures listed in this study and [2] apply this hydrologically based classification system on constraining hydrologic models and predicting streamflow in ungauged basins. The physical descriptors mentioned in this study are only suggestions based on previous studies of similar signatures or from the understanding of what the signature represents. Once physical and climatic characteristics are related to these signatures, a model output constraint system can be created in improve the predictive capability of models independent of the model parameters. See Yadav et al (2007) for a more complete discussion of the usefulness of this technique.

Bibliography

Abrahams, A. D. (1984). Channel networks: a geomorphological perspective. *Water Resources Research.* 20 (2). 161–188.

Andrews, D. F. (1972). Plots of High-Dimensional Data. *Biometrics,* 28 (1). 125-136. doi:10.2307/2528964.

Arnold, J. G., Allen, P. M. (1999). Automated methods for estimating baseflow and ground water recharge from streamflow records. *Journal of the American Water Resources Association.* 35 (2). 411-424.

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

Beighley, R. E., Dunne, T., and Melack, J. M. (2005). Understanding and modeling basin hydrology: interpreting the hydrogeological signature. *Hydrological Processes*. 19. 1333-1353.

Berne, A., Uijlenhoet, R., and Troch, P. A. (2005). Similarity analysis of subsurface flow response of hillslopes with complex geometry. *Water Resources Research.* 41. doi:10.1029/2004WR003629.

Beven, K. J. (2000). On the uniqueness of place and process representations in hydrological modelling. *Hydrology and Earth System Science*. 4 (2). 203–212.

Beven, K. J. and Kirkby, M. J. (1979). A physically based variable contributing area model of basin hydrology. *Hydrologic Science Bulletin.* 24 303-325.

Black, P. E. (1997). Watershed functions. *Journal of the American Water Resources Association.* 33. 1-11.

Bloeschl, G. and Sivapalan, M. (1995). Scale issues in hydrological modeling: A review. *Hydrological Processes.* 9. 251-290.

Bras, R. (1990). *Hydrology: an introduction to hydrologic science.* Addison-Wesley Publishing Company. Reading, MA.

Brutsaert, W. and Nieber, J. L. (1977). Regionalized Drought Flow Hydrographs From a Mature Glaciated Plateau. *Water Resources Research*. 13. 637-643.

Clapp, R. B. and Hornberger, G. M. (1978). Empirical equations for some soil hydraulic properties. *Water Resources Research*. 14. 601-604.

Clark, M. P. and Vrugt, J. A. (2006). Unraveling uncertainties in hydrologic model calibration: addressing the problem of compensatory parameters. *Geophysical Research Letters*. 33. L06406.

Clark, M. P., Slater, A. G., Barrett, A. P., Hay, L. E., McCabe, G. J., Rajagopalan, B., and Leavesley, G. H. (2006). Assimilation of snow covered area information into hydrologic and land-surface models. *Advances in Water Resources*. 29. 1209-1221.

Clausen, B., and Biggs, B. J. F. (2000). Flow variables for ecological studies in temperate streams: grouping based on covariance. *Journal of Hydrology.* 237. 184–197.

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. (2006). Model Parameter Estimation Experiment (MOPEX): Overview and Summary of the Second and Third Workshop Results. *Journal of Hydrology.* 320. 3-17.

Eckhardt, K. (2005). How to construct recursive digital filters for Baseflow separation. *Hydrologic Processes*. 19. 507-515.

Eckhardt, K. (2007). A comparison of baseflow indices, which were calculated with seven different baseflow separation methods. *Journal of Hydrology*. 352. 168-173.

Falkenmark, M., Rockstrom, J., and Savenije, H. (2004). *Balancing water for humans and nature: the new approach in ecohydrology*. Earthscan. London.

Figueiredo, M. A. T, Jain, A. K. (2002). Unsupervised learning of finite mixture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 24. 381-396.

Fu, G. Charles, S. P., and 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. W11419.

Gordon, N. D., McMahon, T. A., Finlayson, B. L., Gippel, C. J., and Nathan, R. J. (2004). *Stream Hydrology - An Introduction for Ecologists Edition Two*. John Wiley & Sons Ltd. West Sussex, England.

Haines, A. T., Finlayson, B. L., and McMahon, T. A. (1988). A global classification of river regimes. *Applied Geography.* 8. 255–272.

Horton, R. E. (1945). Erosional development of streams and their drainage basins; hydrophysical approach to quantitative morphology. *Bulletin of the Geological Society of America.* 56. 275-370.

Institute of Hydrology. (1980). Low-flow studies. *Reort No 1*, Wallingford, Oxon, U.K.

Jensco, K. G., McGlynn, B. L., Gooseff, M. N., Wondzell, S. M., Bencala, K. E, and Marshall, L. A. (2009). Hydrologic connectivity between landscapes and streams: Transferring reach- and plot-scale understanding to the catchment scale. *Water Resources Research.* 45. W04428

Kroll, C., Luz, J., Allen, B., Vogel, R. M. (2004). Devoloping a Watershed Characteristics Database to Improve Low Streamflow Prediction. *Journal of Hydrologic Engineering*. 2. 116-125.

Laaha, G., Bloschl, G. (2006). A comparison of low flow regionalization methods – catchment grouping. *Journal of Hydrology.* 323. 193-214.

Lagacherie, P., Cazemier, D. R., van Gaans, P. F. M., and Burrough, P. A. (1997). Fuzzy k-means clustering of fields in an elementary catchment and extrapolation to a larger area. Geoderma. 77. 197-216.

Langbein W.B. (1947).Topographic characters of drainage basins. *United States Geological Survey. Water-Supply Paper*. 968-C. 125-158.

Lim, K.J., Engel, B. A., Tang, Z., Choi, J., Kim, K., Muthukrishnan, S., Tripathy, D. (2005). Automated Web GIS Based Hydrograph Analysis Tool, WHAT. *JAWRA*. 04133. 1407-1416.

Lyne V. D., Hollick M. (1979). Stochastic time-variable rainfall runoff modelling. *Hydrology and Water Resources Symposium*, Institution of Engineers Australia, Perth; 89–92.

McDonnell, J. J., and Woods, R. (2004). On the need for catchment classification. *Journal of Hydrology.* 299. 2–3.

Merz, R. and Bloschl, G. (2005). Flood frequency regionalisation – spatial proximity vs. catchment attributes. *Journal of Hydrology.* 302. 283-306.

Milly, P. C. D. (1994). Climate, soil water storage, and the average water balance. *Water Resources Research.* 30. 2143-2156.

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

O'Sullivan, D., Unwin, D. J. (2003). *Geographic Information Analysis*. Wiley & Sons Inc. Hoboken, NJ.

Olden, J. D., and Poff, N. L. (2003). Redundancy and the choice of hydrologic indices for characterizing streamflow regimes. *River Research and Applications.* 19. 101–121.

Pike, J. G. (1964). The estimation of annual runoff from meteorological data in a tropical climate. *Journal of Hydrology.* 2. 116–123.

Poff, N., Olden, J. D., Pepin, D. M., Bledsoe, B. P. (2006). Placing global stream flow variability in geographic and geomorphic contexts. *River Research and Applications.* 22. 149–166.

Richter, B. D., Baumgartner, J. V., Powell, J., and Braun, D. P. (1996). A method for assessing hydrologic alteration within ecosystems. *Conservation Biology.* 10. 1163–1174.

Sankarasubramanian, A., Vogel, R. M. (2003). Hydroclimatology of the continental United States. *Geophysical Research Letters.* 30. No. 7, 1363

Sankarasubramanian, A., Vogel, R. M., Limbrunner, J. F. (2001). Climate elasticity of streamflow in the United States. *Water Resources Research*. 377. 1771-1781

Schumm, S.A. (1956). Evolution of drainage systems and slopes in badlands at Perth Amboy, New Jersey. *Geological Society of America Bulletin*. 67. 597-646.

Schwartz, M. D., Ahas, R., and Aasa, A. (2006). Onset of spring starting earlier across the northern hemisphere. *Global Change Biology*. 12. 343-351.

Sivapalan, M. (2005). Pattern, processes and function: elements of a unified theory of hydrology at the catchment scale. In: Anderson, M. (ed.) Encyclopedia of hydrological sciences. London, John Wiley, pp. 193–219.

Strahler, A. N. (1957). Quantitative analysis of watershed geomorphology. *TransActions of the American Geophysical Union.* 38 (6). 913 – 920.

Tillotson, P. M., and Nielsen, D. R. (1984). Scale factors in soil science. *Soil Science Society of America Journal.* 48. 953-959.

Torfs, P. and Wojcik, R. (2001). Local probabilistic neural networks in hydrology. *Physics and Chemistry of the Earth.* 26. 9-14.

Van Haveren, B. P. (1986). Water Resource Measurements: A Handbook for Hydrologists and Engineers. *American Water Works Assn*. Denver, CO.

Vogel, R. M., Kroll, C. N. (1992). Regional Geohydrologic-Geomorphic Relationships for the Estimation of Low-Flow Statistics. *Water Resources Research*. 28. 2451-2458.

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

Wagener, T., Sivapalan, M., Troch, P., and Woods, R. (2007). Catchment Classification and Hydrologic Similarity. *Geography Compass* 1/4 2007. 901-931.

Ward Jr., J. H. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association.* 58. 236-244.

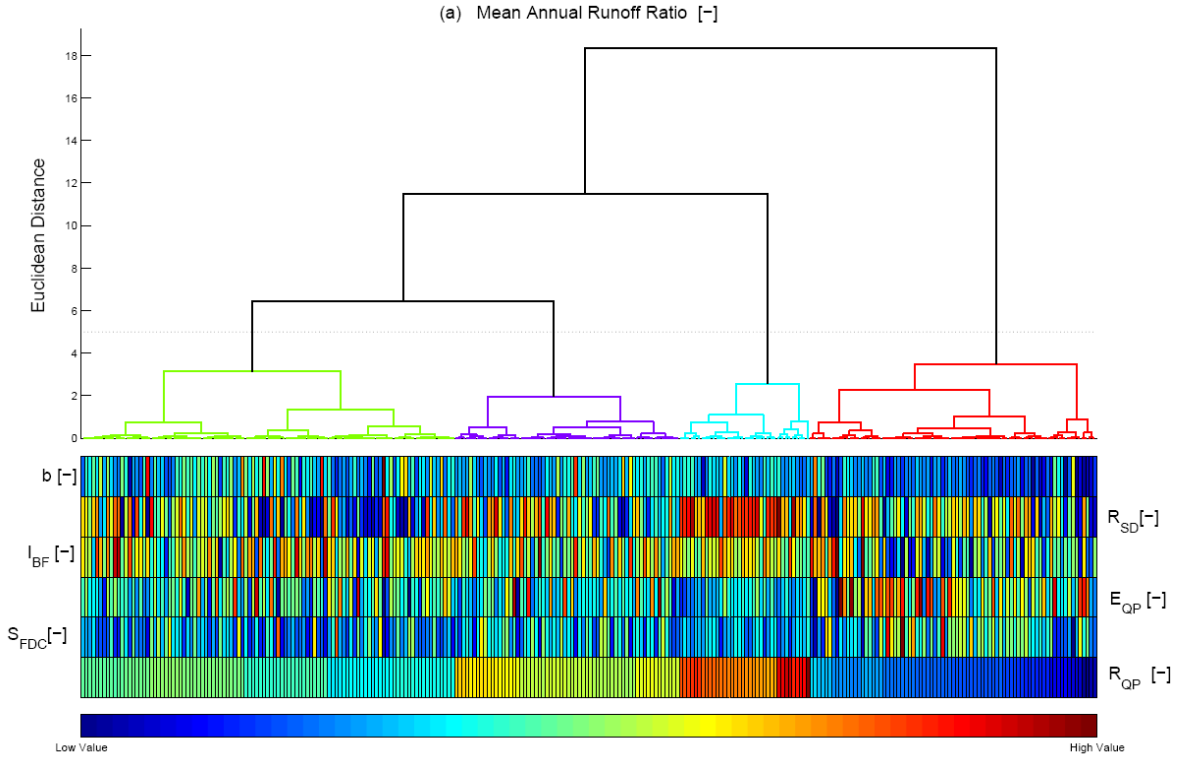
Weiskel P. K., Vogel, R. M., Steeves, P. A., Zarriello, P. J., DeSimone, L. A., and Ries III, K. G. (2007). Water use regimes: Characterizing direct human interaction with hydrologic systems. *Water Resources Research.* 43. doi:10.1029/2006WR005062.

Wittenberg, H. and Sivapalan, M. (1999). Watershed groundwater balance estimation using streamflow recession analysis ad baseflow separation. *Journal of Hydrology*. 219. 20-33.

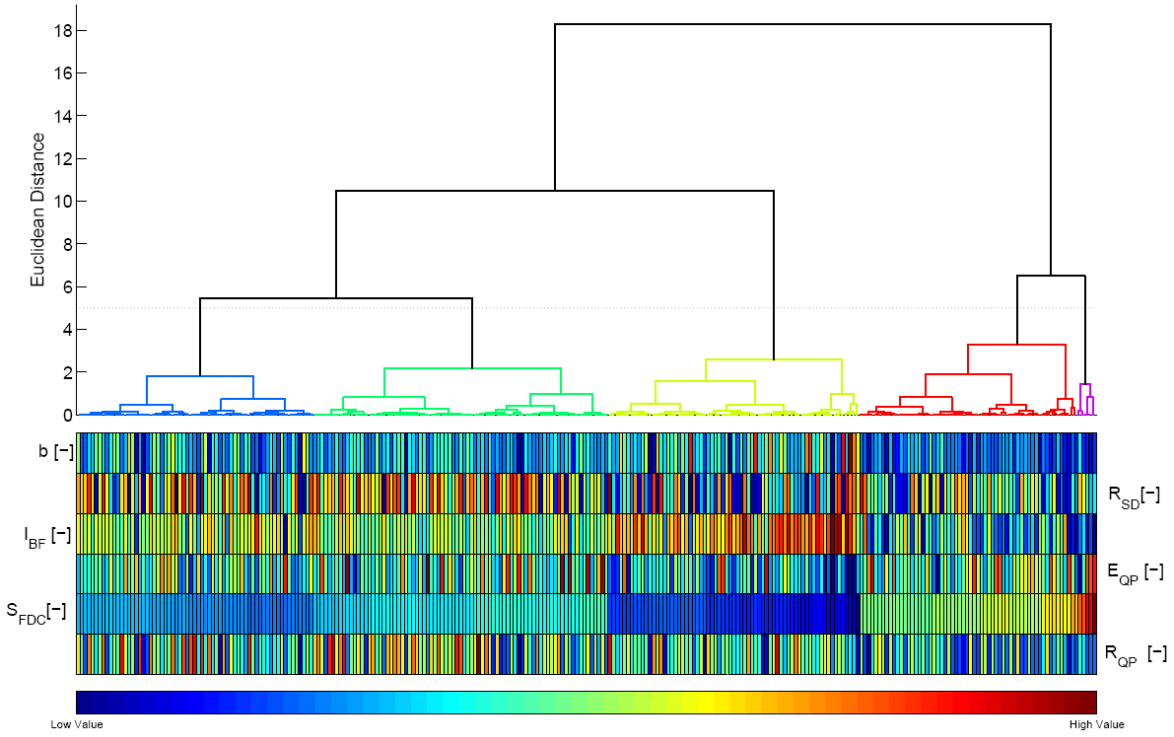
Woods, R. A., (2003). The relative roles of climate, soil, vegetation and topography in determining seasonal and long-term catchment dynamics. *Advances in Water Resources.* 26 (3). 295–309.

Yadav, M., Wagener, T., and Gupta, H. (2007). Regionalization of constraints on expected watershed Response. *Advances in Water Resources.* 30. 1756-1774.

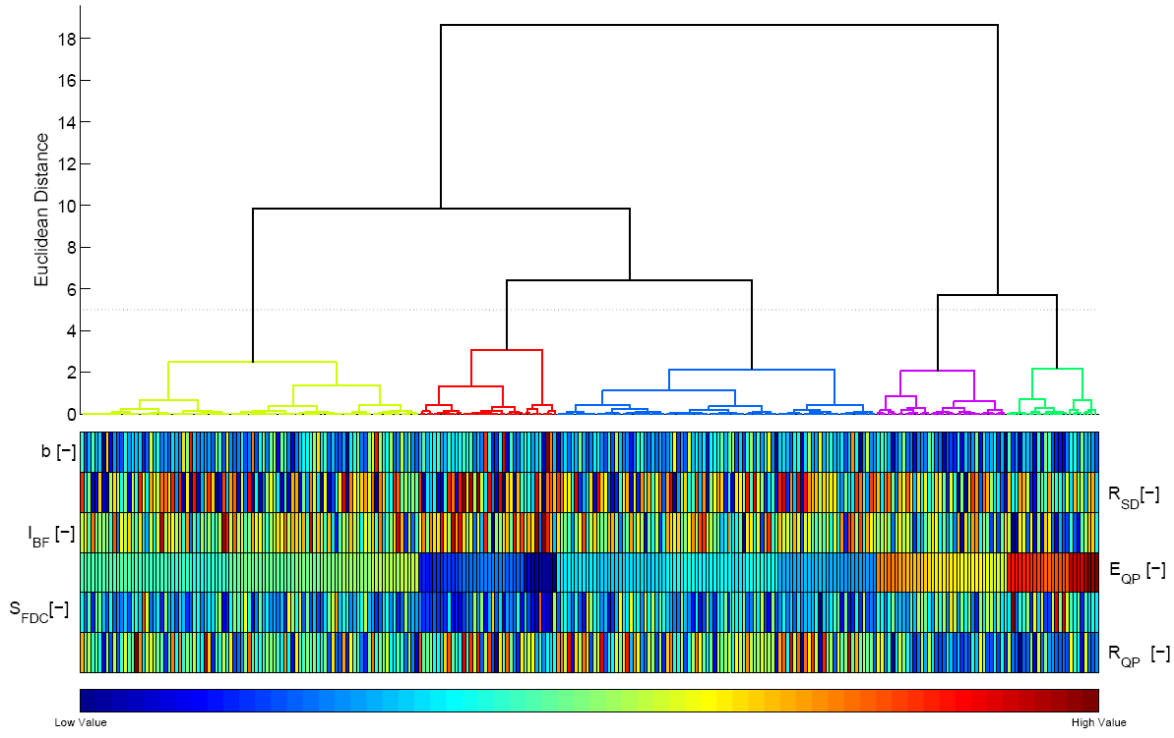
Appendix


(a) Mean Annual Runoff Ratio [−]
Euclidean Distance
18
16
14
12
10
8
6
4
2
0
b [−]
I_{BF} [−]
S_{FDC}[−]
R_{SD}[−]
E_{QP} [−]
R_{QP} [−]
Low Value
High Value

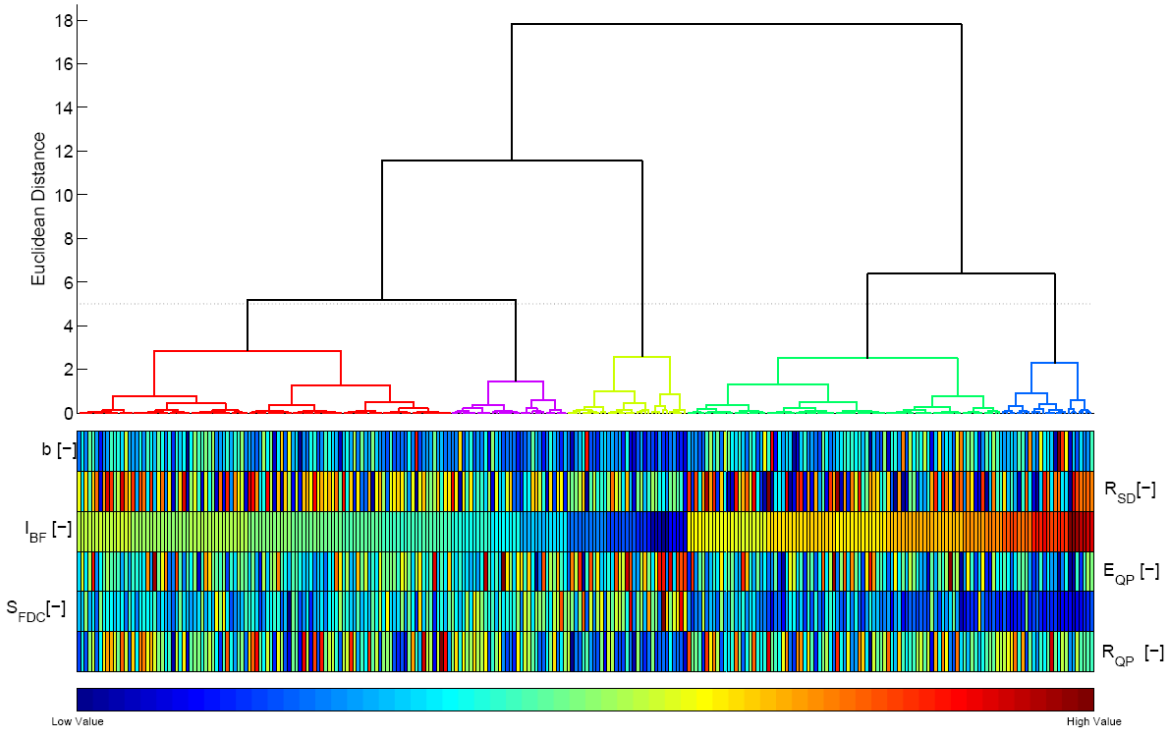
(b) Slope of Flow Duration Curve [−]



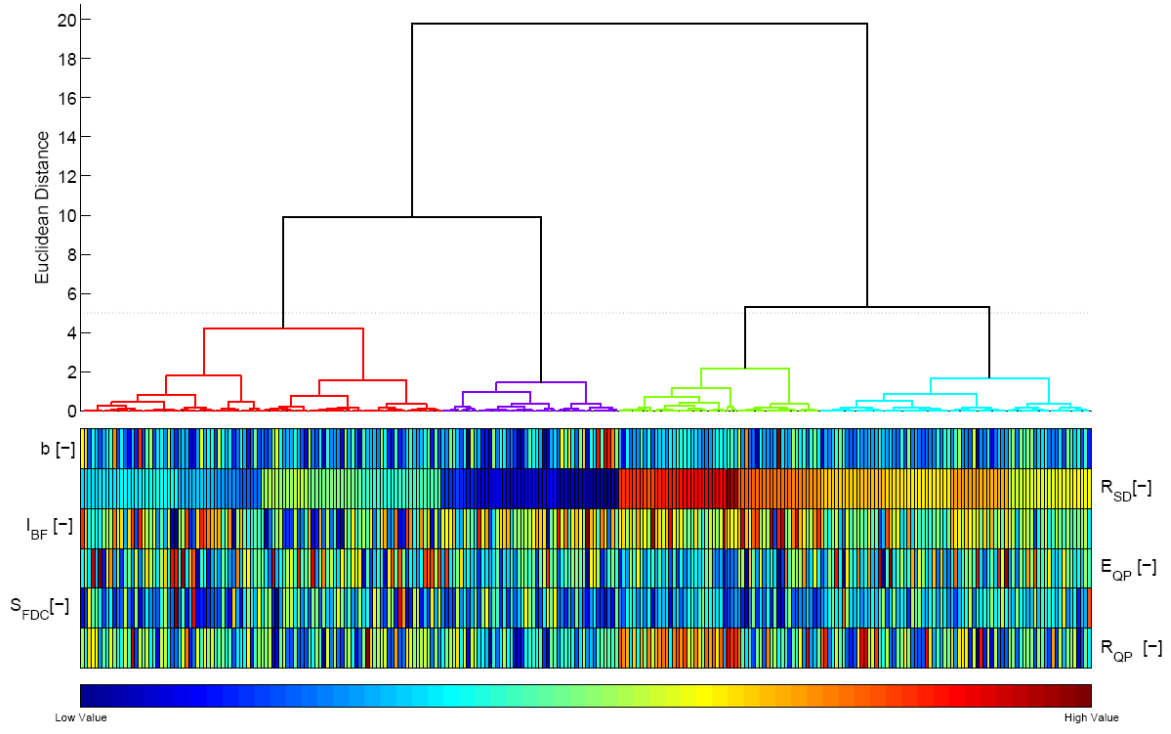
(c) Annual Streamflow Elasticity [−]



(d) Base Flow Index [−]



(e) Ratio of Snow Days [−]



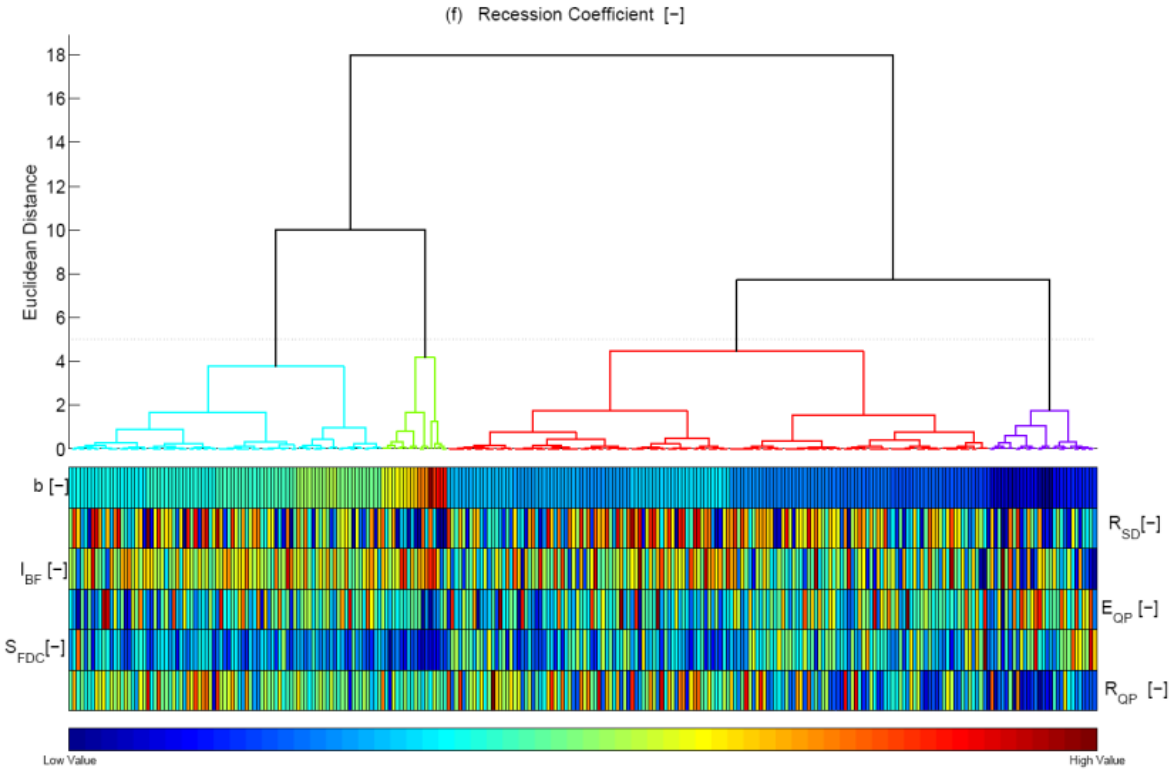


Figure A-1. Recursive pattern plot ordered using dendrogram of individual signature hierarchical cluster analysis. Each leaf within the dendrogram represents a catchment signature value and each row in the recursive pattern plot represents a signature. The signatures represented include, from bottom to top, Runoff Ratio, Slope of FDC, Streamflow Elasticity, Base Flow Index, Snow Day Ratio, and Recession Coefficient.

Table A-1. Summary of all catchments considered for analysis. Contains the name of the river/channel/stream and the town the gauge is located in or near, latitude of the gauge, longitude of the gauge, and the district code pertaining to the catchment.

Site #	Station Name	Latitude [°]	Longitude [°]	District Code
01055500	Nezinscot River at Turner Center, ME	44.269	-70.230	23
01064500	Saco River near Conway, NH	43.991	-71.091	23
01076500	Pemigewasset River at Plymouth, NH	43.759	-71.686	33
01127000	Quinebaug River at Jewett City, CT	41.598	-71.984	9

01197500	Housatonic River near Great Barrington, MS	42.232	-73.355	25
01200000	Tenmile River near Gaylordsville, CT	41.659	-73.528	9
01321000	Sacandaga River near Hope, NY	43.353	-74.270	36
01334500	Hoosic River near Eagle Bridge, NY	42.939	-73.377	36
01348000	East Canada Creek at East Creek, NY	43.017	-74.741	36
01371500	Wallkill River at Gardiner, NY	41.686	-74.165	36
01372500	Wappinger Creek near Wappingers Falls, NY	41.653	-73.873	36
01421000	East Br Delaware R at Fishs Eddy, NY	41.973	-75.174	36
01423000	West Branch Delaware River at Walton, NY	42.166	-75.140	36
01426500	West Branch Delaware River at Hale Eddy, NY	42.003	-75.384	36
01445500	Pequest River at Pequest, NJ	40.831	-74.978	34
01500500	Susquehanna River at Unadilla, NY	42.321	-75.317	36
01503000	Susquehanna River at Conklin, NY	42.035	-75.803	36
01512500	Chenango River near Chenango Forks, NY	42.218	-75.848	36
01518000	Tioga River at Tioga, PA	41.908	-77.129	42
01520000	Cowanesque River near Lawrenceville, PA	41.997	-77.140	42
01520500	Tioga River at Lindley, NY	42.029	-77.132	36
01531000	Chemung River at Chemung, NY	42.002	-76.635	36
01534000	Tunkhannock Creek near Tunkhannock, PA	41.558	-75.895	42
01541000	West Branch Susquehanna River at Bower, PA	40.897	-78.677	42
01541500	Clearfield Creek at Dimeling, PA	40.972	-78.406	42
01543500	Sinnemahoning Creek at Sinnemahoning, PA	41.317	-78.103	42
01548500	Pine Creek at Cedar Run, PA	41.522	-77.447	42
01556000	Frankstown Br Juniata River at Williamsburg, PA	40.463	-78.200	42

01558000	Little Juniata River at Spruce Creek, PA	40.613	-78.141	42
01559000	Juniata River at Huntingdon, PA	40.485	-78.019	42
01560000	Dunning Creek at Belden, PA	40.072	-78.493	42
01562000	Raystown Branch Juniata River at Saxton, PA	40.216	-78.265	42
01567000	Juniata River at Newport, PA	40.478	-77.129	42
01574000	West Conewago Creek near Manchester, PA	40.082	-76.720	42
01595000	North Branch Potomac River at Steyer, MD	39.302	-79.307	24
01606500	So. Branch Potomac River near Petersburg, WV	38.991	-79.176	54
01608500	South Branch Potomac River near Springfield, WV	39.447	-78.654	54
01610000	Potomac River at Paw Paw, WV	39.539	-78.456	24
01611500	Cacapon River near Great Cacapon, WV	39.582	-78.310	54
01628500	S F Shenandoah River near Lynnwood, VA	38.323	-78.755	51
01631000	S F Shenandoah River at Front Royal, VA	38.914	-78.211	51
01634000	N F Shenandoah River near Strasburg, VA	38.977	-78.336	51
01643000	Monocacy River at Jug Bridge Near Frederick, MD	39.403	-77.366	24
01649500	North East Branch Anacostia River at Riverdale, MD	38.960	-76.926	24
01663500	Hazel River at Rixeyville, VA	38.592	-77.965	51
01664000	Rappahannock River at Remington, VA	38.531	-77.814	51
01667500	Rapidan River near Culpeper, VA	38.350	-77.975	51
01668000	Rappahannock River near Fredericksburg, VA	38.308	-77.529	51
01672500	South Anna River near Ashland, VA	37.797	-77.549	51
01674500	Mattaponi River near Beulahville, VA	37.884	-77.165	51
02016000	Cowpasture River near Clifton Forge, VA	37.792	-79.759	51
02018000	Craig Creek at Parr, VA	37.666	-79.911	51

02030500	Slate River near Arvonia, VA	37.703	-78.378	51
02055000	Roanoke River at Roanoke, VA	37.258	-79.939	51
02083500	Tar River at Tarboro, NC	35.894	-77.533	37
02102000	Deep River at Moncure, NC	35.627	-79.116	37
02116500	Yadkin River at Yadkin College, NC	35.857	-80.387	37
02118000	South Yadkin River near Mocksville, NC	35.845	-80.659	37
02126000	Rocky River near Norwood, NC	35.149	-80.176	37
02135000	Little Pee Dee R. at Galivants Ferry, SC	34.057	-79.247	45
02138500	Linville River near Nebo, NC	35.795	-81.890	37
02143000	Henry Fork near Henry River, NC	35.684	-81.403	37
02143040	Jacob Fork at Ramsey, NC	35.591	-81.567	37
02143500	Indian Creek near Laboratory, NC	35.421	-81.265	37
02156500	Broad River near Carlisle, SC	34.595	-81.421	45
02165000	Reedy River near Ware Shoals, SC	34.417	-82.151	45
02192000	Broad River near Bell, GA	33.974	-82.770	13
02202500	Ogeechee River near Eden, GA	32.192	-81.416	13
02217500	Middle Oconee River near Athens, GA	33.947	-83.423	13
02228000	Satilla River at Atkinson, GA	31.221	-81.866	13
02236000	St. Johns River near De Land, FL	29.008	-81.383	125
02273000	Kissimmee River at S-65e Near Okeechobee, FL	27.226	-80.963	125
02296750	Peace River at Arcadia, FL	27.221	-81.876	124
02329000	Ochlockonee River near Havana, FL	30.554	-84.384	123
02339500	Chattahoochee River at West Point, GA	32.886	-85.182	13
02347500	Flint River at Us 19, near Carsonville, GA	32.721	-84.233	13

02349500	Flint River at Montezuma, GA	32.298	-84.044	13
02365500	Choctawhatchee River At Caryville, FL	30.776	-85.828	123
02375500	Escambia River near Century, FL	30.965	-87.234	123
02383500	Coosawattee River near Pine Chapel, GA	34.564	-84.833	13
02387000	Conasauga River at Tilton, GA	34.667	-84.928	13
02387500	Oostanaula River at Resaca, GA	34.577	-84.942	13
02414500	Tallapoosa River at Wadley, AL	33.117	-85.561	1
02448000	Noxubee River at Macon, MS	33.102	-88.562	28
02456500	Locust Fork at Sayre, AL	33.710	-86.983	1
02472000	Leaf River near Collins, MS	31.707	-89.407	28
02475000	Leaf River near Mclain, MS	31.103	-88.808	28
02475500	Chunky River near Chunky, MS	32.326	-88.909	28
02478500	Chickasawhay River at Leakesville, MS	31.149	-88.548	28
02479300	Red Creek at Vestry, MS	30.736	-88.781	28
02482000	Pearl River at Edinburg, MS	32.799	-89.335	28
02486000	Pearl River at Jackson, MS	32.281	-90.179	28
02492000	Bogue Chitto River near Bush, LA	30.629	-89.897	22
03010500	Allegheny River at Eldred, PA	41.963	-78.386	42
03011020	Allegheny River at Salamanca, NY	42.156	-78.715	36
03020500	Oil Creek at Rouseville, PA	41.482	-79.695	42
03024000	French Creek at Utica, PA	41.438	-79.956	42
03032500	Redbank Creek at St. Charles, PA	40.995	-79.394	42
03050500	Tygart Valley River near Elkins, WV	38.924	-79.879	54
03051000	Tygart Valley River at Belington, WV	39.029	-79.936	54

03054500	Tygart Valley River at Philippi, WV	39.150	-80.039	54
03065000	Dry Fork at Hendricks, WV	39.072	-79.623	54
03069000	Shavers Fork at Parsons, WV	39.096	-79.677	54
03069500	Cheat River near Parsons, WV	39.123	-79.681	54
03070000	Cheat River at Rowlesburg, WV	39.346	-79.665	54
03075500	Youghiogheny River near Oakland, MD	39.422	-79.424	24
03079000	Casselman River at Markleton, PA	39.860	-79.228	42
03109500	Little Beaver Creek near East Liverpool, OH	40.676	-80.541	39
03111500	Short Creek near Dillonvale, OH	40.193	-80.734	39
03114500	Middle Island Creek at Little, WV	39.475	-80.997	54
03155500	Hughes River at Cisco, WV	39.119	-81.277	54
03159500	Hocking River at Athens, OH	39.329	-82.088	39
03161000	South Fork New River near Jefferson, NC	36.393	-81.407	37
03164000	New River near Galax, VA	36.647	-80.979	51
03167000	Reed Creek at Grahams Forge, VA	36.939	-80.887	51
03168000	New River at Allisonia, VA	36.938	-80.746	51
03173000	Walker Creek at Bane, VA	37.268	-80.710	51
03175500	Wolf Creek near Narrows, VA	37.306	-80.850	51
03179000	Bluestone River near Pipestem, WV	37.544	-81.010	54
03180500	Greenbrier River at Durbin, WV	38.544	-79.833	54
03182500	Greenbrier River at Buckeye, WV	38.186	-80.131	54
03183500	Greenbrier River at Alderson, WV	37.724	-80.641	54
03184000	Greenbrier River at Hilldale, WV	37.640	-80.805	54
03186500	Williams River at Dyer, WV	38.379	-80.484	54

03198500	Big Coal River at Ashford, WV	38.180	-81.712	54
03237500	Ohio Brush Creek near West Union OH	38.804	-83.421	39
03238500	White Oak Creek near Georgetown OH	38.858	-83.929	39
03251500	Licking River at Mckinneysburg, KY	38.598	-84.267	21
03252500	South Fork Licking River at Cynthiana, KY	38.391	-84.303	21
03253500	Licking River at Catawba, KY	38.710	-84.311	21
03265000	Stillwater River at Pleasant Hill OH	40.058	-84.356	39
03266000	Stillwater River at Englewood OH	39.870	-84.286	39
03269500	Mad River near Springfield OH	39.923	-83.870	39
03274000	Great Miami River at Hamilton OH	39.391	-84.572	39
03281500	South Fork Kentucky River at Booneville, KY	37.480	-83.675	21
03289500	Elkhorn Creek near Frankfort, KY	38.269	-84.815	21
03301500	Rolling Fork near Boston, KY	37.767	-85.704	21
03303000	Blue River near White Cloud, IN	38.238	-86.228	18
03308500	Green River at Munfordville, KY	37.268	-85.886	21
03324300	Salamonie River near Warren, IN	40.713	-85.454	18
03326500	Mississinewa River at Marion, IN	40.576	-85.660	18
03328500	Eel River near Logansport, IN	40.782	-86.264	18
03331500	Tippecanoe River near Ora, IN	41.157	-86.564	18
03339500	Sugar Creek at Crawfordsville, IN	40.049	-86.899	18
03345500	Embarras River at Ste. Marie, IL	38.936	-88.023	17
03346000	North Fork Embarras River near Oblong, IL	39.010	-87.946	17
03348000	White River at Anderson, IN	40.106	-85.672	18
03349000	White River at Noblesville, IN	40.047	-86.017	18

03361500	Big Blue River at Shelbyville, IN	39.529	-85.782	18
03362500	Sugar Creek near Edinburgh, IN	39.361	-85.997	18
03363000	Driftwood River near Edinburgh, IN	39.339	-85.986	18
03364000	East Fork White River at Columbus, IN	39.200	-85.926	18
03365500	East Fork White River at Seymour, IN	38.983	-85.899	18
03381500	Little Wabash River at Carmi, IL	38.061	-88.160	17
03410500	South Fork Cumberland River near Stearns, KY	36.627	-84.533	21
03438000	Little River near Cadiz, KY	36.778	-87.722	21
03443000	French Broad River at Blantyre, NC	35.299	-82.624	37
03451500	French Broad River at Asheville, NC	35.609	-82.578	37
03455000	French Broad River near Newport, TN	35.982	-83.161	47
03465500	Nolichucky River at Embreeville, TN	36.176	-82.457	47
03473000	S F Holston River near Damascus, VA	36.652	-81.844	51
03504000	Nantahala River near Rainbow Springs, NC	35.128	-83.619	37
03512000	Oconaluftee River at Birdtown, NC	35.461	-83.354	37
03521500	Clinch River at Richlands, VA	37.086	-81.781	51
03524000	Clinch River at Cleveland, VA	36.945	-82.155	51
03528000	Clinch River above Tazewell, TN	36.425	-83.398	47
03531500	Powell River near Jonesville, VA	36.662	-83.095	51
03532000	Powell River near Arthur, TN	36.542	-83.630	47
03540500	Emory River at Oakdale, TN	35.983	-84.558	47
03550000	Valley River at Tomotla, NC	35.139	-83.981	37
03574500	Paint Rock River near Woodville AL	34.624	-86.306	1
04073500	Fox River at Berlin, WI	43.954	-88.953	55

04079000	Wolf River at New London, WI	44.392	-88.740	55
04100500	Elkhart River at Goshen, IN	41.593	-85.849	18
04113000	Grand River at Lansing, MI	42.751	-84.555	26
04115000	Maple River at Maple Rapids, MI	43.110	-84.693	26
04165500	Clinton River at Moravian Drive At Mt. Clemens, MI	42.596	-82.909	26
04176500	River Raisin near Monroe, MI	41.961	-83.531	26
04178000	St. Joseph River near Newville, IN	41.385	-84.802	18
04185000	Tiffin River at Stryker OH	41.504	-84.430	39
04191500	Auglaize River near Defiance OH	41.238	-84.399	39
04198000	Sandusky River near Fremont OH	41.308	-83.159	39
04201500	Rocky River near Berea OH	41.407	-81.887	39
04223000	Genesee River at Portageville NY	42.570	-78.042	36
05053000	Wild Rice River near Abercrombie, ND	46.468	-96.784	38
05244000	Crow Wing River at Nimrod, MN	46.640	-94.879	27
05280000	Crow River at Rockford, MN	45.087	-93.734	27
05320500	Le Sueur River near Rapidan, MN	44.111	-94.041	27
05408000	Kickapoo River at La Farge, WI	43.574	-90.643	55
05410490	Kickapoo River at Steuben, WI	43.183	-90.858	55
05412500	Turkey River at Garber, IA	42.740	-91.262	19
05418500	Maquoketa River near Maquoketa, IA	42.083	-90.633	19
05422000	Wapsipinicon River near De Witt, IA	41.767	-90.535	19
05430500	Rock River at Afton, WI	42.609	-89.071	55
05435500	Pecatonica River at Freeport, IL	42.303	-89.620	17
05440000	Kishwaukee River near Perryville, IL	42.194	-88.999	17

05447500	Green River near Geneseo, IL	41.489	-90.158	17
05451500	Iowa River at Marshalltown, IA	42.066	-92.908	19
05452000	Salt Creek near Elberon, IA	41.964	-92.313	19
05454500	Iowa River at Iowa City, IA	41.657	-91.541	19
05455500	English River at Kalona, IA	41.470	-91.715	19
05458500	Cedar River at Janesville, IA	42.648	-92.465	19
05462000	Shell Rock River at Shell Rock, IA	42.712	-92.583	19
05471500	South Skunk River near Oskaloosa, IA	41.356	-92.657	19
05472500	North Skunk River near Sigourney, IA	41.301	-92.205	19
05476500	Des Moines River at Estherville, IA	43.397	-94.844	19
05479000	East Fork Des Moines River at Dakota City, IA	42.724	-94.193	19
05481000	Boone River near Webster City, IA	42.432	-93.806	19
05482500	North Raccoon River near Jefferson, IA	41.988	-94.377	19
05484500	Raccoon River at Van Meter, IA	41.534	-93.950	19
05508000	Salt River near New London, MO	39.612	-91.407	29
05514500	Cuivre River near Troy, MO	39.009	-90.978	29
05515500	Kankakee River at Davis, IN	41.400	-86.701	18
05517000	Yellow River at Knox, IN	41.303	-86.621	18
05517500	Kankakee River at Dunns Bridge, IN	41.220	-86.968	18
05518000	Kankakee River at Shelby, IN	41.183	-87.340	18
05520500	Kankakee River at Momence, IL	41.160	-87.669	17
05526000	Iroquois River near Chebanse, IL	41.009	-87.823	17
05542000	Mazon River near Coal City, IL	41.286	-88.360	17
05552500	Fox River at Dayton, IL	41.384	-88.789	17

05554500	Vermilion River at Pontiac, IL	40.878	-88.636	17
05569500	Spoon River at London Mills, IL	40.708	-90.280	17
05570000	Spoon River at Seville, IL	40.490	-90.340	17
05582000	Salt Creek near Greenview, IL	40.132	-89.736	17
05584500	La Moine River at Colmar, IL	40.330	-90.896	17
05585000	La Moine River at Ripley, IL	40.025	-90.632	17
05592500	Kaskaskia River at Vandalia, IL	38.961	-89.089	17
05593000	Kaskaskia River at Carlyle, IL	38.612	-89.356	17
05594000	Shoal Creek near Breese, IL	38.610	-89.495	17
06441500	Bad R near Fort Pierre, SD	44.327	-100.384	46
06480000	Big Sioux River near Brookings, SD	44.180	-96.749	46
06600500	Floyd River at James, IA	42.577	-96.311	19
06606600	Little Sioux River at Correctionville, IA	42.471	-95.798	19
06607200	Maple River at Mapleton, IA	42.157	-95.810	19
06609500	Boyer River at Logan, IA	41.642	-95.783	19
06799500	Logan Creek near Uehling, NE	41.713	-96.522	31
06808500	West Nishnabotna River at Randolph, IA	40.873	-95.580	19
06810000	Nishnabotna River above Hamburg, IA	40.633	-95.626	19
06811500	Little Nemaha River at Auburn, NE	40.393	-95.813	31
06813000	Tarkio River at Fairfax, MO	40.339	-95.406	29
06815000	Big Nemaha River at Falls City, NE	40.036	-95.596	31
06817000	Nodaway River at Clarinda, IA	40.739	-95.013	19
06820500	Platte River near Agency, MO	39.688	-94.703	29
06847000	Beaver Creek near Beaver City, NE	40.120	-99.893	31

06860000	Smoky Hill R at Elkader, KS	38.795	-100.858	20
06869500	Saline R at Tescott, KS	39.004	-97.874	20
06883000	Little Blue River near Deweese, NE	40.333	-98.067	31
06884400	L Blue R near Barnes, KS	39.726	-96.805	20
06885500	Black Vermillion R near Frankfort, KS	39.682	-96.443	20
06888500	Mill C near Paxico, KS	39.065	-96.169	20
06892000	Stranger C near Tonganoxie, KS	39.116	-95.011	20
06894000	Little Blue River near Lake City, MO	39.101	-94.301	29
06897500	Grand River near Gallatin, MO	39.927	-93.943	29
06898000	Thompson River at Davis City, IA	40.640	-93.808	19
06899500	Thompson River at Trenton, MO	40.069	-93.638	29
06908000	Blackwater River at Blue Lick, MO	38.992	-93.197	29
06913500	Marais Des Cygnes R near Ottawa, KS	38.618	-95.268	20
06914000	Pottawatomie C near Garnett, KS	38.334	-95.249	20
06933500	Gasconade River at Jerome, MO	37.930	-91.977	29
07019000	Meramec River near Eureka, MO	38.506	-90.592	29
07029500	Hatchie River at Bolivar, Tn	35.275	-88.977	47
07052500	James River at Galena, MO	36.805	-93.462	29
07056000	Buffalo River near St. Joe, AR	35.983	-92.747	5
07057500	North Fork River near Tecumseh, MO	36.623	-92.248	29
07058000	Bryant Creek near Tecumseh, MO	36.627	-92.306	29
07067000	Current River at Van Buren, MO	36.991	-91.014	29
07068000	Current River at Doniphan, MO	36.622	-90.848	29
07069500	Spring River at Imboden, AR	36.205	-91.172	5

07072000	Eleven Point River near Ravenden Springs, AR	36.347	-91.114	5
07074000	Strawberry River near Poughkeepsie, AR	36.110	-91.450	5
07144200	L Arkansas River at Valley Center, KS	37.832	-97.389	20
07147070	Whitewater River at Towanda, KS	37.796	-97.014	20
07147800	Walnut River at Winfield, KS	37.224	-96.996	20
07152000	Chikaskia River near Blackwell, OK	36.811	-97.277	40
07163000	Council Creek near Stillwater, OK	36.116	-96.868	40
07172000	Caney River near Elgin, Ks	37.004	-96.317	20
07177500	Bird Creek near Sperry, OK	36.278	-95.954	40
07183000	Neosho River near Iola, Ks	37.922	-95.428	20
07186000	Spring River near Waco, MO	37.246	-94.566	29
07196500	Illinois River near Tahlequah, OK	35.923	-94.924	40
07197000	Baron Fork at Eldon, OK	35.921	-94.839	40
07243500	Deep Fork near Beggs, OK	35.674	-96.069	40
07252000	Mulberry River near Mulberry. AR	35.577	-94.016	5
07261000	Cadron Creek near Guy, AR	35.299	-92.404	5
07288500	Big Sunflower River at Sunflower, MS	33.547	-90.543	28
07290000	Big Black River nr Bovina, MS	32.348	-90.697	28
07340000	Little River near Horatio, AR	33.919	-94.387	5
07346070	Little Cypress Ck near Jefferson, TX	32.713	-94.346	48
07348000	Twelvemile Bayou near Dixie, LA	32.646	-93.878	22
07363500	Saline River near Rye, AR	33.701	-92.026	5
07375500	Tangipahoa River at Robert, LA	30.507	-90.362	22
07378000	Comite River near Comite, LA	30.513	-91.074	22

07378500	Amite River near Denham Springs, LA	30.464	-90.990	22
08013500	Calcasieu River near Oberlin, LA	30.640	-92.814	22
08015500	Calcasieu River near Kinder, LA	30.503	-92.915	22
08032000	Neches River near Neches, TX	31.892	-95.431	48
08033500	Neches River near Rockland, TX	31.025	-94.399	48
08055500	Elm Fk Trinity River near Carrollton, TX	32.966	-96.944	48
08085500	Clear Fk Brazos River at Ft Griffin, TX	32.935	-99.225	48
08095000	N Bosque River nr Clifton, TX	31.786	-97.568	48
08146000	San Saba River at San Saba, TX	31.213	-98.719	48
08150000	Llano River near Junction, TX	30.504	-99.735	48
08167500	Guadalupe River near Spring Branch, TX	29.860	-98.384	48
08171000	Blanco River at Wimberley, TX	29.994	-98.089	48
08171300	Blanco River near Kyle, TX	29.979	-97.910	48
08172000	San Marcos River at Luling, TX	29.666	-97.651	48
08189500	Mission River at Refugio, TX	28.292	-97.279	48
08205500	Frio River near Derby, TX	28.737	-99.145	48

Table A-2. Summary of all catchments considered for analysis. Includes catchment attributes of area, stream gauge elevation, and the porosity of the dominant soil.

Site #	Area	Stream gauge elevation	Porosity of Dominant Soil
01055500	169	276.29	0.4100
01064500	385	418.19	0.4350
01076500	622	457.07	0.4100
01127000	713	63.07	0.4350
01197500	282	683.04	0.4350
01200000	203	304	0.4510
01321000	491	881.31	0.4350
01334500	510	355.41	0

01348000	289	335.7	0.4350
01371500	695	185.7	0.4850
01372500	181	114.37	0.4850
01421000	784	955.96	0
01423000	332	1190.3	0.4850
01426500	595	946.46	0.4850
01445500	106	398.78	0.4350
01500500	982	997.25	0.4850
01503000	2232	841.04	0.4850
01512500	1483	871.63	0.4850
01518000	282	1021.07	0.4850
01520000	298	983.96	0.4850
01520500	771	964.5	0.4850
01531000	2506	778.63	0.4850
01534000	383	610.1	0.4510
01541000	315	1207.14	0.4510
01541500	371	1146.08	0.4510
01543500	685	769.36	0.4510
01548500	604	780.36	0.4510
01556000	291	831.78	0.4510
01558000	220	751.15	0.4510
01559000	816	599.69	0.4350
01560000	172	1051.16	0.4850
01562000	756	795.77	0.4510
01567000	3354	363.93	0.4510
01574000	510	263.68	0.4850
01595000	73.1	2276.01	0
01606500	651	967.87	0
01608500	1461	561.41	0
01610000	3129	487.88	0
01611500	675	456.78	0
01628500	1079	1013.17	0.4820
01631000	1634	469.38	0
01634000	770	494.03	0
01643000	817	231.92	0.4850
01649500	72.8	12.68	0.4350
01663500	285	288.3	0.4760
01664000	619	252.53	0.4760
01667500	468	241.36	0.4760
01668000	1595	70	0.4760
01672500	395	83.74	0.4820
01674500	603	12.43	0.4200
02016000	461	1006.93	0

02018000	329	992.5	0
02030500	226	238.78	0.4820
02055000	384	906.84	0.4820
02083500	2183	9.32	0.4820
02102000	1434	185.06	0.4770
02116500	2280	638.45	0.4820
02118000	306	663.62	0.4820
02126000	1372	212.91	0.4770
02135000	2790	23.95	0.4200
02138500	66.7	1203.87	0.4510
02143000	83.2	890.99	0.4350
02143040	25.7	1103	0.4760
02143500	69.2	741.62	0.4820
02156500	2790	290.79	0.4820
02165000	236	451.14	0.4820
02192000	1430	357.19	0.4820
02202500	2650	17.64	0.4200
02217500	398	555.66	0.4820
02228000	2790	14.79	0.4200
02236000	3066	-	0.3950
02273000	-	-	0.3950
02296750	1367	6	0.3950
02329000	1140	59.36	0.4200
02339500	3550	551.67	0.4820
02347500	1850	334.54	0.4350
02349500	2900	255.83	0.4350
02365500	3499	39.02	0.4200
02375500	3817	28.34	0.4200
02383500	831	616.16	0
02387000	687	622.28	0
02387500	1602	604.14	0
02414500	1675	599.87	0.4820
02448000	768	142	0.4820
02456500	885	258.64	0
02472000	743	197.01	0.4820
02475000	3495	42.15	0.4350
02475500	369	269	0.4510
02478500	2690	51.13	0.4350
02479300	441	20.1	0.4350
02482000	904	341.67	0.4820
02486000	3171	233.7	0.4850
02492000	1213	44.25	0.4510
03010500	550	1416.53	0.4850

03011020	1608	1358	0.4850
03020500	283	1028.32	0.4510
03024000	1028	1019.44	0.4850
03032500	528	973.14	0.4510
03050500	271	1894.5	0
03051000	406	1679.62	0
03054500	914	1279.96	0
03065000	349	1699.26	0
03069000	213	1634.87	0
03069500	722	1589.66	0
03070000	939	1368.24	0
03075500	134	2353.61	0
03079000	382	1655.29	0.4510
03109500	496	702.77	0.4510
03111500	123	676.1	0.4770
03114500	458	631.32	0.4510
03155500	453	607.92	0
03159500	943	611.26	0.4770
03161000	205	2657.04	0.4510
03164000	1141	2208.04	0.4510
03167000	258	1924.65	0.4820
03168000	2212	1848.36	0.4510
03173000	299	1665.92	0.4820
03175500	223	1583.83	0
03179000	395	1526.91	0.4510
03180500	133	2699.5	0
03182500	540	2085.69	0
03183500	1364	1529.01	0
03184000	1619	1388.31	0
03186500	128	2194.17	0
03198500	391	621.83	0.4350
03237500	387	510.6	0.4770
03238500	218	604.2	0.4770
03251500	2326	520.83	0
03252500	621	688.52	0
03253500	3300	500.01	0
03265000	503	846.73	0.4510
03266000	650	780	0.4510
03269500	490	881.42	0.4510
03274000	3630	499.98	0.4510
03281500	722	642.49	0.4770
03289500	473	540.2	0.4920
03301500	1299	400.42	0

03303000	476	434.26	0.4850
03308500	1673	451.7	0.4820
03324300	425	784.65	0.4770
03326500	682	774.56	0.4770
03328500	789	621.5	0.4770
03331500	856	692.91	0.4510
03339500	509	657.77	0.4770
03345500	1516	445.75	0.4770
03346000	318	456.19	0.4770
03348000	406	825.02	0.4510
03349000	858	738.16	0.4510
03361500	421	737.67	0.4510
03362500	474	646.23	0.4510
03363000	1060	636.99	0.4510
03364000	1707	603.12	0.4510
03365500	2341	550.67	0.4510
03381500	3102	339.91	0.4850
03410500	954	764.81	0.4510
03438000	244	391.45	0.4850
03443000	296	2060.32	0.4510
03451500	945	1950.28	0.4510
03455000	1858	1011.61	0.4510
03465500	805	1519.3	0.4510
03473000	303	1792.3	0.4510
03504000	51.9	3072.97	0.4510
03512000	184	1843.3	0.4510
03521500	137	1924.08	0.4820
03524000	533	1500.24	0.4820
03528000	1474	1060.7	0.4820
03531500	319	1259.08	0.4510
03532000	685	1043.84	0.4510
03540500	764	761.38	0
03550000	104	1556.46	0.4510
03574500	320	570.95	0.4820
04073500	1340	744.52	0.3950
04079000	2260	747.94	0.4350
04100500	594	769.43	0.4510
04113000	1230	805.53	0.4510
04115000	434	642.58	0.4510
04165500	734	570.43	0.4510
04176500	1042	616.26	0.4770
04178000	610	795.4	0.4770
04185000	410	685.1	0.4770

04191500	2318	659.7	0.4770
04198000	1251	626.3	0.4770
04201500	267	649.9	0.4770
04223000	984	1080	0.4850
05053000	2080	907.94	0.4510
05244000	1030	1313.27	0.3950
05280000	2640	893.08	0.4510
05320500	1110	775.76	0.4760
05408000	266	781.54	0.4850
05410490	687	657	0.4850
05412500	1545	634.46	0.4510
05418500	1553	625.96	0.4510
05422000	2336	598.81	0.4510
05430500	3340	742.14	0.4770
05435500	1326	743.18	0.4770
05440000	1099	692.13	0.4770
05447500	1003	577.66	0.4850
05451500	1532	853.1	0.4510
05452000	201	781.58	0.4770
05454500	3271	617.27	0.4770
05455500	574	633.45	0.4770
05458500	1661	868.26	0.4510
05462000	1746	885.34	0.4510
05471500	1635	685.5	0.4510
05472500	730	651.53	0.4770
05476500	1372	1247.55	0.4510
05479000	1308	1038.71	0.4510
05481000	844	989.57	0.4510
05482500	1619	967.09	0.4510
05484500	3441	841.16	0.4510
05508000	2480	477.03	0.4770
05514500	903	450.27	0.4770
05515500	537	664.68	0.3950
05517000	435	679.93	0.4510
05517500	1352	649.65	0.3950
05518000	1779	628.13	0.3950
05520500	2294	609.18	0.3950
05526000	2091	595.99	0.4770
05542000	455	527.41	0.4770
05552500	2642	462.3	0.4770
05554500	579	619.45	0.4770
05569500	1072	508.97	0.4770
05570000	1636	467.04	0.4770

05582000	1804	479	0.4770
05584500	655	491.53	0.4770
05585000	1293	431.1	0.4770
05592500	1940	453.3	0.4770
05593000	2719	402.92	0.4770
05594000	735	413.97	0.4770
06441500	3107	1427.83	0
06480000	3898	1551.91	0.4850
06600500	886	1092.59	0.4770
06606600	2500	1096.49	0.4510
06607200	669	1085.86	0.4850
06609500	871	1009.38	0.4850
06799500	1015	1208.73	0.4850
06808500	1326	932.99	0.4770
06810000	2806	894.17	0.4770
06811500	792	889.87	0.4770
06813000	508	867.66	0.4770
06815000	1339	858.24	0.4760
06817000	762	955.36	0.4770
06820500	1760	807.38	0.4770
06847000	2080	2164.84	0.4850
06860000	3555	2622.62	0.4850
06869500	2820	1265.34	0.4850
06883000	984	1632.67	0.4850
06884400	3351	1140.06	0.4770
06885500	410	1106.91	0.4820
06888500	318	964.92	0.4920
06892000	406	800.95	0.4760
06894000	184	719.15	0.4770
06897500	2250	707.56	0.4760
06898000	701	874.04	0.4770
06899500	1720	721.87	0.4760
06908000	1120	593.79	0.4770
06913500	1250	857.68	0.4920
06914000	334	873.23	0.4920
06933500	2840	657.64	0.4770
07019000	3788	404.18	0.4770
07029500	1480	323.49	0.4760
07052500	987	921.37	0.4770
07056000	829	560.35	0.4920
07057500	561	584.67	0.4850
07058000	570	573.15	0.4850
07067000	1667	442.78	0.4770

07068000	2038	321.21	0.4770
07069500	1180	254.07	0.4820
07072000	1130	291.98	0.4820
07074000	473	298.07	0.4820
07144200	1327	1325.66	0.4920
07147070	426	1231.47	0.4920
07147800	1880	1082.86	0.4920
07152000	1859	967.41	0.4920
07163000	31	828.28	0.4760
07172000	445	763.32	0
07177500	905	579.43	0
07183000	3723	928.92	0.4920
07186000	1164	833.23	0.4850
07196500	959	664.14	0.4850
07197000	307	701.14	0.4850
07243500	2018	632.55	0.4350
07252000	373	432.75	0.4920
07261000	169	371.68	0
07288500	767	93	0.4770
07290000	2812	84.93	0.4850
07340000	2660	300	0
07346070	675	174.6	0.4200
07348000	3137	136.12	0.4200
07363500	2100	97.06	0.4850
07375500	646	6.6	0.4850
07378000	284	23.85	0.4850
07378500	1280	-	0.4850
08013500	753	39.43	0.4850
08015500	1700	11.95	0.4850
08032000	1145	264.06	0.3950
08033500	3636	88.41	0.4350
08055500	2459	431.4	0.4820
08085500	3988	1174	0.4820
08095000	968	605.43	0
08146000	3046	1162.16	0
08150000	1854	1634.32	0
08167500	1315	948.1	0
08171000	355	797.23	0.4760
08171300	412	620.12	0
08172000	838	322.05	0.4820
08189500	690	-	0.4200
08205500	3429	449.11	0.4760
**Information taken from USGS			

Table A-3. Summary of the 45 possible signatures used, including units and a brief description.

Name	Units	Description
Mean Annual Q	[m]	Amount of mean annual streamflow
Mean Value Daily Jan Q	[mm]	Mean value of January streamflow
Mean Value Daily Feb Q	[mm]	Mean value of February streamflow
Mean Value Daily March Q	[mm]	Mean value of March streamflow
Mean Value Daily April Q	[mm]	Mean value of April streamflow
Mean Value Daily May Q	[mm]	Mean value of May streamflow
Mean Value Daily June Q	[mm]	Mean value of June streamflow
Mean Value Daily July Q	[mm]	Mean value of July streamflow
Mean Value Daily Aug Q	[mm]	Mean value of August streamflow
Mean Value Daily Sept Q	[mm]	Mean value of September streamflow
Mean Value Daily Oct Q	[mm]	Mean value of October streamflow
Mean Value Daily Nov Q	[mm]	Mean value of November streamflow
Mean Value Daily Dec Q	[mm]	Mean value of December streamflow
Variance in Jan Q	[mm^2]	Interannual variance of mean January streamflow
Variance in Feb Q	[mm^2]	Interannual variance of mean February streamflow
Variance in March Q	[mm^2]	Interannual variance of mean March streamflow
Variance in April Q	[mm^2]	Interannual variance of mean April streamflow
Variance in May Q	[mm^2]	Interannual variance of mean May streamflow
Variance in June Q	[mm^2]	Interannual variance of mean June streamflow
Variance in July Q	[mm^2]	Interannual variance of mean July streamflow
Variance in Aug Q	[mm^2]	Interannual variance of mean August streamflow

Variance in Sept Q	[mm^2]	Interannual variance of mean September streamflow
Variance in Oct Q	[mm^2]	Interannual variance of mean October streamflow
Variance in Nov Q	[mm^2]	Interannual variance of mean November streamflow
Variance in Dec Q	[mm^2]	Interannual variance of mean December streamflow
High Pulse Count 3	[-]	Number of days per year that streamflow exceeds 3 times the median streamflow value
High Pulse Count 3.5	[-]	Number of days per year that streamflow exceeds 3.5 times the median streamflow value
High Pulse Count 4	[-]	Number of days per year that streamflow exceeds 4 times the median streamflow value
High Pulse Count 4.5	[-]	Number of days per year that streamflow exceeds 4.5 times the median streamflow value
High Pulse Count 5	[-]	Number of days per year that streamflow exceeds 5 times the median streamflow value
High Pulse Count 5.5	[-]	Number of days per year that streamflow exceeds 5.5 times the median streamflow value
High Pulse Count 6	[-]	Number of days per year that streamflow exceeds 6 times the median streamflow value
High Pulse Count 6.5	[-]	Number of days per year that streamflow exceeds 6.5 times the median streamflow value
High Pulse Count 7	[-]	Number of days per year that streamflow exceeds 7 times the median streamflow value
Flow Value 75%	[mm]	75th percentile streamflow value
High Pulse Count 75%	[-]	Number of days per year above the 75th percentile streamflow value
Recession Intercept, a	[d^{-1}] x b**	a coefficient in the relationship, dQ/dt=aQ^b
SlopeFDC_Q/ SlopeFDC_P	[-]	Slope of the flow duration curve divded by the slope of the precipitation duration curve at the 33rd and 66th percentiles
Monthly Streamflow Elasticity	[-]	Streamflow elasticity calculated at the inter-monthly time scale
Mean Annual Runoff Ratio ***	[-]	Mean annual streamflow devided by the mean annual precipitation
Annual Streamflow Elasticity ***	[-]	Streamflow elasticity calculated at the inter-annual time scale
Base Flow Index ***	[-]	Amount of baseflow contribution to streamflow divided by the amount of total streamflow
Slope of Flow Duration Curve ***	[-]	Slope of the streamflow duration curve between 33rd and 66th percentile exceedence values normalized by the mean streamflow

Ratio of Snow Days ***	[-]	Number of days with precipitation when mean temperature is below 2 degrees C divided by the number of days with precipitation
Recession Coefficient, b ***	[-]	b coefficient in the relationship, dQ/dt=aQ^b
**Units dependent on the value of the Reseccion Coefficient. See Wittenberg and Sivapalan (1999) ***Key signatures used in analysis		