DISSERTATION

IMPROVING HYDROLOGIC MODELING OF UNGAGED BASINS TO SUPPORT ENVIRONMENTAL FLOW MANAGEMENT IN A HETEROGENEOUS REGION

Submitted by

Stephen K. Adams

Department of Civil and Environmental Engineering

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2021

Doctoral Committee:

Advisor: Brian P. Bledsoe

N. LeRoy Poff
Jefferey D. Niemann
Eric D. Stein

ABSTRACT

IMPROVING HYDROLOGIC MODELING OF UNGAGED BASINS TO SUPPORT ENVIRONMENTAL FLOW MANAGEMENT IN A HETEROGENEOUS REGION

Environmental streamflow management can sustain aquatic ecosystems and the services they provide by reestablishing elements of the natural flow regime that are necessary for ecological health. One of the more difficult challenges with developing environmental flow criteria is estimating streamflow at locations without gage data; however, this challenge is not unique to environmental flows. Streamflow prediction in ungaged basins is a very common problem in hydrology and engineering with no clear solution, but it is particularly difficult to model environmental streamflow metrics across heterogeneous regions with highly diverse land uses, geologic settings, and hydroclimatological processes.

In this dissertation, I create a new regionalization framework, “Streamflow Regionalization with Hydrologic Model-based Classification” (SR-HMC), for modeling challenging flow metrics in ungaged basins across a heterogeneous region. I also test the efficacy of the new framework for developing environmental streamflow criteria. In Chapter 2, I explore different approaches for classifying streams with similar flow regimes and develop a novel classification technique for prioritizing regional accuracy of hydrologic models. As the precursor to SR-HMC, this “Hydrologic Model-based Classification” (HMC) groups hydrologically similar streams by determining the degree of reciprocity of calibrated parameters between a regional catalog of rainfall-runoff models as quantified through jackknife resampling. Results show that HMC complements traditional classifications based on streamflow metrics and watershed characteristics, and offer advantages over these traditional classifications when used to regionalize ungaged basins.

Next, Chapter 3 describes implementation of ensemble modeling to optimize HMC into a regionalization framework for producing time series of streamflow at ungaged sites. For gaged locations, hydrologic model parameters that cannot be calculated directly can be calibrated using observed flows; however, these same model parameters are much more uncertain and difficult to estimate at ungaged locations. SR-HMC uses geographically-weighted model output averaging with regionally-calibrated parameter sets to reduce parameter uncertainty in models of ungaged basins. This new framework is tested at five sites across a large and diverse region. Results were improved using SR-HMC over standard nearest-neighbor regionalization approaches.

Finally, I turn to management applications of these novel methods in ungaged basins by analyzing the statistical relationships between streamflow alteration and ecological integrity. In Chapter 4, I compare the explanatory power of simple flow-ecology relationships produced by different methods for regionalizing ungaged basins and different metrics of flow alteration. Results highlight robust modeling practices amenable to management. Development of environmental streamflow recommendations based on prediction in ungaged basins is an ongoing challenge; however, this research demonstrates how novel approaches to classification and model extrapolation can improve streamflow estimation at ungaged locations in heterogeneous regions, and thereby bolster the scientific basis of environmental flow management.

ACKNOWLEDGEMENTS

Foremost, I would like to thank my advisor, Brian Bledsoe, for his endless support and patience over my time at CSU. Despite moving to Georgia, Brian has always made time for me. His guidance and advice have greatly helped my professional and personal development. He has been instrumental in helping me get to where I am in life. Thank you for all your wisdom.

To the rest of my committee, Erin Stein, LeRoy Poff, and Jeff Niemann, thank you for your support, guidance, and inspiration. Eric and the team at SCCWRP, especially Rafi Mazor, Ashmita Sengupta, and Kenny McCune, were instrumental in obtaining funding for this work from the California State Water Resources Control Board, and were closely involved with model development. I would also like to thank Phil Turk from the CSU Statistics Department for inspiring my work with jackknife resampling.

Thank you to my colleagues and friends from CSU, especially Rod Lammers, Sarah Parker, Tyler Rosburg, Joel Sholtes, Johannes Beeby, Dan Baker, Tim Stephens, Travis Stroth, Travis Hardee, Erin Ryan, Dan Brogan, and Jacob Morgan, for all the advice and good times we shared. I would also like to thank my colleagues at Tetra Tech for inspiring me and showing me how to be a better engineer.

Finally, thank you to my family and close friends for their support over the years. To my sister Elizabeth, especially, your resolve motivates me. I could not have done it without all of you.

DEDICATION

To my family, friends, and colleagues who helped me along my way, especially my advisor Brian Bledsoe.

And to the CSU community- past, present, and future. You inspire me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

Human activities and land use changes are altering streamflows and stressing freshwater ecosystems on a global scale (Allan, 2004; Bunn and Arthington, 2002; Richter *et al.*, 1997). Environmental streamflow management can mitigate degradation of freshwater ecosystems by defining the key environmental flow elements that are necessary for sustaining these systems (Poff *et al*., 1997). Environmental flows describe magnitudes, frequencies, durations, timing, and rates of change needed to support aquatic ecosystems, and the economic, cultural, and human health services they provide to society (Arthington *et al*., 2018; Poff *et al*., 1997). Healthy ecosystems offer familiar and palpable services including recreation, water supply, and aesthetic value, but they also offer significant and quantifiable non-tangible benefits through conservation, stewardship, and responsible management (Wilson and Carpenter, 1999). When ecological degradation is caused by hydrologic modification, environmental streamflow management can be used as a tool to reestablish the natural flow regime for restoring ecological services. Specifically, the methods of the Ecological Limits of Flow Alteration (ELOHA) framework have been used to support regional environmental flow development via the natural flow regime (Poff *et al*., 2010).

ELOHA is a consensus framework for determining the components of the natural flow regime most critical to ecological integrity across an entire region. ELOHA facilitates recovery of the natural flow regime by developing quantitative relationships between flow alterations and changes in biological conditions. It implements a variety of methods to support environmental flow management, which can be broken down into four steps: 1) develop a hydrologic foundation of the altered and “baseline” natural flow regimes; 2) classify streams from the region into groups with distinct flow regimes; 3) quantify deviation of streamflow between baseline and altered conditions; and 4) develop flow-ecology relationships using estimates of flow alteration and ecological integrity. These flow-ecology relationships provide the tool for water managers to incorporate specific environmental flows into regional management. For a given site, paired streamflow and biological data are needed for pre- and post-altered hydrological conditions to fully develop flow-ecology relationships describing how biological communities respond to streamflow changes (Poff *et al.*, 2010; Poff and Zimmerman, 2010).

Bioassessment sites often occur on small, wadable streams with insufficient gage data (Poff *et al*., 2006), which necessitates hydrologic modeling of ungaged basins. Estimating streamflow at ungaged locations is a challenge commonly encountered not only during ELOHA, but in many other analyses of streams and watersheds. Established techniques for modeling ungaged basins are recommended by ELOHA and have been studied extensively, yet there is no consensus approach and parameter uncertainty is a major challenge (Blöschl *et al.*, 2013). This uncertainty is increased when modeling ungaged basins across a heterogeneous region, where diverse land use, geological settings, and hydroclimatological processes contribute to a wide range of flow regimes among geographically nearby streams (Arsenault and Brissette, 2014; Arsenault *et al*., 2019; Blöschl *et al.*, 2013). Furthermore, environmental flow metrics are often volatile and subsequently challenging to model (Blöschl *et al*., 2013; Nathan and McMahon, 1992; Razavi and Coulibaly, 2017). As such, each of ELOHA's four steps offer opportunities to develop, implement, and assess new methods for modeling ungaged basins within the context of environmental streamflows. My dissertation develops, applies, and tests new statistical and modeling approaches to reduce parameter uncertainty in hydrologic models of ungaged basins across a heterogeneous region, pertinent not only to ELOHA, but also other scientific, engineering, and management applications.

ELOHA has been applied all over the world, from Australia (Mackay, 2014), to Spain (Solans and García de Jalón, 2016), to China (Zhang *et al*., 2012). In the United States, ELOHA has been explored to address the effects of urbanization on stream biota across the country from the Pacific Northwest (Cassin *et al.*, 2005), to the Atlantic coast (Buchanan *et al.*, 2013; Kennen *et al.*, 2013; McManamay *et al*., 2012; Pomeroy *et al.*, 2008), and many other places. My dissertation studies coastal southern California (So. CA) where the ELOHA framework has also been implemented, and where water resources management has always been a tremendous challenge. Few regions of the country have experienced more rapid and extensive urbanization over the past few decades than So. CA, including the San Diego and Los Angeles metropolitan areas (Hawley and Bledsoe, 2011). In So. CA, the combination of a Mediterranean and semi-arid climate, explosive population growth, rapid land development, and tremendous agricultural production have strained water resources and significantly altered natural streamflows (Stein *et al*., 2012). Increases in impervious surfaces associated with urbanization are a particularly potent and widespread driver of flow alteration that decrease infiltration and amplify runoff (Arnold and Gibbons, 1996), increase hydrograph flashiness (Walsh *et al.*, 2005), alter stream geomorphology (Paul and Meyer, 2001), decrease water quality (Allan, 2004), and increase sediment transport capacity (Hawley and Bledsoe, 2013). These effects of human development are exacerbated for freshwater systems in semi-arid regions like So. CA (Falkenmark *et al*., 1989). Previous work with ELOHA in So. CA has involved preliminary hydrologic foundations (Parker *et al.*, 2019; Sengupta *et al.*, 2018) and flow-ecology relationships (Mazor *et al.*, 2018; Parker *et al.*, 2019), state-wide stream classification (Pyne *et al.*, 2017), a recommendation of preliminary flow targets (Stein *et al.*, 2017a), and an implementation of those targets in the San Diego River watershed (Stein *et al.*, 2017b). These previous efforts have required hydrologic modeling of ungaged basins, which has been challenging in the highly heterogeneous region of So. CA characterized by broad gradients in land use, geologic settings, hydroclimatological processes, and streamflow regimes.

This dissertation builds on previous work in environmental streamflows and modeling ungaged basins to bolster the scientific foundation of ELOHA. Ultimately, I develop a new method for modeling ungaged basins that reduces parameter uncertainty. I achieve this by first grouping hydrologically-similar streams in called “Hydrologic Model-based Classification” (HMC). I then expand HMC into a full framework for modeling ungaged basins in a heterogeneous region, named “Streamflow Regionalization with Hydrologic Model-based Classification” (SR-HMC). SR-HMC is applied to develop the hydrologic foundation of ELOHA in So. CA by estimating environmental streamflow metrics traditionally challenging to model. The hydrologic foundation produced by SR-HMC is compared to common nearest neighbor regionalization methods for modeling ungaged basins. This comparison of methods for regionalizing ungaged basins is carried throughout the entire ELOHA framework, as levels of further analyses are added to stream classification, quantifying hydrologic alteration, and formulating flow-ecology relationships. While the tools and ideas proposed in my dissertation have been developed through the lens of ELOHA and environmental streamflows, my novel methods to reduce parameter uncertainty in ungaged basins can be applied to any study of regional streamflows. Specific objectives within the dissertation are provided in each subsequent chapter.

In Chapter 2, I classify regional streams at gage locations in So. CA based on hydrologic and physical watershed similarity. Furthermore, I develop “Hydrologic Model-based Classification” to prioritize the accuracy of hydrologic models in ungaged basins. HMC is ideal when regional ungaged basins must be modeled, especially in a heterogeneous region, and it complements other methods of grouping streams with similar flow regimes.

In Chapter 3, I apply ensemble modeling to HMC to create “Streamflow Regionalization with Hydrologic Model-based Classification” for modeling ungaged basins. SR-HMC combines statistical, hydrologic, and ensemble modeling to quantify and reduce parameter uncertainty from a catalog of regional calibrated parameter sets. It is applicable to analyses of ungaged basins outside the development of environmental flows but is most appropriate for diverse regions where neighboring basins may differ hydrologically.

In Chapter 4, I formulate flow-ecology relationships at five ungaged sites while comparing different methods for quantifying flow alteration and regionalizing ungaged basins. SR-HMC is assessed alongside nearest neighbor regionalization for estimating flow alteration and flow-ecology relationships in ungaged basins. This final chapter demonstrates how to improve the integrity of flow-ecology relationships and confidence in environmental streamflow management in ungaged basins. It ends with specific recommendations of modeling practices in ungaged basins that may assist environmental streamflow management.

Chapter 2

Advancing stream classification and hydrologic modeling of ungaged basins for environmental flow management in coastal southern California

Summary

Environmental streamflow management can improve the ecological health of streams by returning modified flows to more natural conditions. The Ecological Limits of Hydrologic Alteration (ELOHA) framework for developing regional environmental flow criteria has been implemented to reverse hydromodification across the heterogenous region of coastal southern California (So. CA) by focusing on two elements of the flow regime: streamflow permanence and flashiness. Within ELOHA, classification groups streams by hydrologic and geomorphic similarity to stratify flow-ecology relationships. Analogous grouping techniques are used by hydrologic modelers to facilitate streamflow prediction in ungaged basins (PUB) through regionalization. Most watersheds, including those needed for stream classification and environmental flow development, are ungaged. Furthermore, So. CA is a highly heterogeneous region spanning a gradient of urbanization, which presents a challenge for regionalizing ungaged basins. In this study, I develop a novel classification technique for PUB modeling that uses an inductive approach to group regional streams by modeled hydrologic similarity followed by deductively determining class membership with hydrologic model errors and watershed metrics. As a new type of classification, this “Hydrologic Model-based Classification” (HMC) prioritizes modeling accuracy, which in turn provides a means to improve model predictions in ungaged basins, while complementing traditional classifications and improving environmental flow management. HMC is developed by calibrating a regional catalog of process-based rainfall-runoff models, quantifying the hydrologic reciprocity of calibrated parameters that would be unknown in ungaged basins, and grouping sites according to hydrologic and physical similarity. HMC was applied to 25 USGS streamflow gages in the south coast region of California and was compared to other hybrid PUB approaches combining inductive and deductive classification. Using an Average Cluster Error metric, results show HMC provided the most hydrologically similar groups according to calibrated parameter reciprocity. Hydrologic Model-based Classification is relatively complex and time-consuming to implement, but it shows potential for advancing ungaged basin management. This study demonstrates the benefits of thorough stream classification using multiple approaches, and suggests that Hydrologic Model-based Classification has advantages for PUB and building the hydrologic foundation for environmental flow management.

Keywords: environmental streamflows, ungaged basins, stream classification, hydrologic modeling, ELOHA, California water management

2.1 Introduction

The natural variability of streamflow regimes, including flow magnitude, duration, frequency, timing, and rate of change (Poff *et al.*, 1997), is crucial for maintaining the ecological integrity of streams (Bunn and Arthington, 2002). Maintenance of aquatic and riparian ecosystem functions is a major priority for water managers; however, streamflow regimes have been altered globally as population growth and development lead to urbanization, dams, flow extraction, and other land use changes (Naiman *et al.*, 1995; Richter *et al.*, 1997). Environmental flow criteria frameworks, such as the Ecological Limits of Flow Alteration (ELOHA) (Poff *et al.*, 2010), are methods for protecting the ecological health of streams from hydrologic alteration by reestablishing essential elements of streamflow and sediment regimes. The ELOHA framework is robust because it synthesizes many flow-ecology relationships from a study area to provide a foundation for developing environmental flow recommendations within an entire municipality or management region (Poff *et al.*, 2010). Such a regional approach has been recommended for the widespread implementation of environmental flows because it allows for effective and comprehensive estimation of environmental streamflow regimes at a wide variety of streams in a large and diverse study area (Arthington *et al.*, 2006). The coastal area of southern California (So. CA) is experiencing substantial hydrologic alteration (Hawley and Bledsoe, 2011) and associated ecological decline (Stein *et al.*, 2012), which has prompted application of ELOHA (Mazor *et al.*, 2018; Parker *et al.*, 2019; Pyne *et al.*, 2017; Sengupta *et al.*, 2018; Stein *et al.*, 2017). The region is highly heterogenous, spanning an extensive range of geology, stream types, and land uses, which presents unique challenges for implementing ELOHA.

Stream classification is one of four major steps within the scientific process of ELOHA used to group hydrologically, or otherwise similar, streams (Poff *et al.*, 2010). Its primary role towards developing environmental flows is to stratify flow-ecology relationships by regional stream type, and to help determine where new bioassessment sites should be placed to strengthen the variety of sites within a region. Olden *et al*. (2012) outlined two overarching approaches to hydrologic classification—those utilizing inductive reasoning (observed or modeled flows) and those utilizing deductive reasoning (watershed data characterizing flow). While the inductive approach benefits from actual measures of discharge, it is often plagued by insufficient gauging networks (Olden *et al*., 2012) and uncertainty modeling ungaged basins (Blöschl *et al.*, 2013). Two mirroring state-wide stream classification studies utilizing both inductive and deductive approaches have recently been performed across California. Pyne *et al.* (2017) first clustered all stream reaches based on similarity of watershed characteristics, then used hydrologic metrics to determine cluster membership and separate reference reaches. Conversely, Lane *et al.* (2017) grouped the natural streamflow regime of all reaches before using watershed characteristics to determine flow type. A third state-wide classification study was performed by Lane *et* *al.* (2018), which unified the classifications of Pyne *et al.* (2017) and Lane *et al.* (2017) by using daily-scale hydrologic baseline archetypes based on dimensionless reference hydrographs. These three stream classification studies focused on characterizing natural flow regimes across California, which is a challenge in the heavily hydrologically modified and heterogeneous Southern Coast hydrologic region of CA (Waananen and Crippen, 1977). Sites from this region did not show strong separate from the rest of CA in previous classifications. While most South Coast streams were classified as “rain and seasonal groundwater” (Lane *et al.*, 2017) or “rain and seasonal groundwater” and “flashy, ephemeral rain” (Lane *et al.*, 2018), not one of the 91 reference gages used to drive the Lane *et al.* (2017) classification fell in the South Coast. Furthermore, streams in the Mohave Desert and Central Valley shared the same “rain and seasonal groundwater” classification and South Coast streams (Lane *et al.*, 2017). Central Valley streams remained grouped with South Coast streams in the unified classification (Lane *et al.*, 2018). Finally, none of the seven classes produced by Pyne *et al.* (2017) were dominated by South Coast streams. The results of these three state-wide classifications indicate developing environmental streamflow criteria for South Coast streams could benefit from a more targeted classification focused on the diverse regional landscape.

Regionalization is a common framework for predicting streamflow in ungaged basins (PUB) that is performed by transferring hydrologic information from gaged systems to ungaged (Blöschl *et al.*, 2013; Razavi and Coulibaly, 2013). While regionalization often employs regression equations to compute singular streamflow metrics, such as peak flow, continuous hydrologic models offer process-based analyses with full hydrograph outputs that can be used to analyze past and future climate, land use, and management scenarios. The application of hydrologic models to these alternative scenarios makes them important for developing the hydrologic foundation within ELOHA (Poff *et al*., 2010). Additionally, a hydrologic foundation often necessitates modeling of ungaged basins because crucial bioassessment sites used to develop flow-ecology relationships often occur on small streams without available representative streamflow data (Poff and Ward, 1989). Despite the clear importance of PUB to ELOHA and other stream management efforts, no superior method for regionalizing hydrologic models has emerged (Blöschl *et al.*, 2013).

In a typical flow regionalization effort with hydrologic models, a network of models is created and calibrated at gaged sites across a study area. For ungaged sites within the network, model parameters that cannot be calculated directly are estimated and/or transferred from the catalog of calibrated models, typically using a measure of spatial proximity, physical similarity, or parameter regression (Oudin *et al.*, 2008; Razavi and Coulibaly, 2013; Samuel *et al.*, 2011). While spatial proximity is generally the preferred regionalization approach (Razavi and Coulibaly, 2013), it is not always superior and is less applicable in highly heterogeneous regions, such as So. CA, where neighboring watersheds may have substantially different geology, land use, and/or climate. These challenges with applying a traditional regionalization approach in a highly heterogenous region provide opportunities for PUB innovations. Furthermore, the technique of grouping similar streams is shared by ELOHA and PUB, which provides an excellent opportunity to explore new approaches for classifying streams with the intention of modeling ungaged basins while developing environmental flow criteria in a highly heterogeneous region.

This study was motivated by a desire to improve the science supporting environmental streamflows in So. CA where flow criteria are under development (Mazor *et al.*, 2018; Parker *et al.*, 2019; Sengupta *et al.*, 2018; Stein *et al.*, 2017). In this study, I develop a new method of stream classification that quantifies hydrologic similarity for regionalizing ungaged basins in a heterogeneous region. I compare this new approach to traditional methods of stream classification using hydrologic and watershed characteristics. Towards this end, this study has three specific objectives:

1) Classify streams in coastal southern California using the best current practices;

2) Develop and implement a new approach for stream classification that prioritizes the accuracy of regional hydrologic models; and
3) Compare the accuracy of traditional classifications versus the new approach for estimating streamflow and flow-ecology relationships in heterogeneous ungaged basins.

I hypothesize that directly incorporating regional model accuracy into a stream classification scheme will provide information complementary to existing deductive and inductive schemes and demonstrate greater ability to accurately model ungaged basins through regionalization, compared to the traditional classifications.

2.2 Methods

2.2.1 Study Area

This study was focused within the large coastal region of southern California, which is roughly bounded by the transverse mountain ranges to the north, Mexico to the south, the peninsular mountain ranges to the east, and the Pacific Ocean to the west. Study watersheds lie within the coastal regions of San Diego, Riverside, Orange, San Bernardino, Los Angeles, Ventura, and Santa Barbara Counties, and are considered within the "South Coast" hydrologic region of CA according to the U.S. Geological Survey (USGS) (Waananen and Crippen, 1977). The climate is characterized as semi-arid and Mediterranean with hot, dry summers and mild, wet winters. Diverse regional topography, geology, and precipitation patterns allow for the natural existence of many stream types, spanning perennial, intermittent, and ephemeral. Land use varies dramatically across the region ranging from heavily urban and suburban sprawl, to significantly agricultural, to rural coastal and mountainous. These diverse land uses profoundly influence streamflows, with particular deviation from natural flow regimes occurring due to the urban centers of Los Angeles and San Diego concurrently with the California State Water Project.

As a first step towards developing environmental flow criteria, only USGS stream gage sites were considered with neighboring bioassessment sites from the California Water Boards' Perennial Streams

Assessment (PSA) within the Surface Water Ambient Monitoring Program (SWAMP). This provided gaged flow estimates at bioassessment sites. Hydrologic surrogacy between gage and bioassessment sites was assumed by ensuring a difference in watershed area of less than 15% with no intervening dams, diversions, reservoirs, or interbasin transfers. Gages from the region were selected to contain high-resolution hourly streamflow data for water years (WY) 2005-2007, which typify relatively wet, average, and dry years consecutively in So. CA (WRCC, 2015). Finally, watersheds of selected gages required sufficient meteorological and landscape data to build minimally calibrated rainfall-runoff models (see 2.2.3.1 Hydrologic Models). An exhaustive search for suitable streamflow records yielded 25 USGS gage sites for classification (Figure 2.1; Table A1).

Figure 2.1: Locations of USGS streamflow gages used for classification.

2.2.2 Traditional Classification

Three types of traditional classification were used in this study: an inductive approach with gaged flow data, a deductive approach utilizing watershed characteristics, and a combined inductive and deductive approach applying both types of data.

2.2.2.1 Inductive Approach

Previous research in So. CA has shown streamflow flashiness and drying have important influence on shaping local benthic macroinvertebrate assemblages (Gasith and Resh, 1999; Mazor *et al*., 2018; Parker *et al*., 2019). This influence makes them strong metrics for developing flow-ecology relationships to guide environmental flow recommendations. To this end, flashiness and drying have been extensively studied for developing regional flow criteria (Mazor *et al.*, 2018; Parker *et al.*, 2019; Pyne *et al.*, 2017; Sengupta *et al.*, 2018; Stein *et al.*, 2017); however, additional elements of the natural flow regime are also important drivers of ecological health in CA (Yarnell *et al.*, 2020). For this study, Richards-Baker Flashiness Index (RBI) (Baker *et al*., 2004) and a metric quantifying the frequency of extremely low flows indicative of drying were computed from the 25 hourly time series of discharge. RBI was calculated according to Equation 2.1, wherein Q_t is the discharge at time t, Q_{t+1} is the discharge at time step after t, and T is the final time step.

Equation 2.1: Richards-Baker Flashiness Index (RBI) (Baker et al., 2004).

$$RBI = \frac{\sum_{t=1}^{T} |Q_{t+1} - Q_t|}{\sum_{t=1}^{T} Q_t}$$

To quantify the frequency of extremely low flows indicative of drying, the fraction of flow record with flow less than 1 cfs was calculated according to Equation 2.2, wherein $N_{Q<1cfs}$ is the number of time steps containing streamflow less than 1 cfs and N is the total number of time steps containing flow data.

Equation 2.2: Fraction of time with flow < 1 cfs.

$$< 1\, cfs = \frac{N_{Q<1cfs}}{N}$$

Although flows less than 1 cfs are recorded by USGS, this threshold was chosen to indicate stream drying given the inherent measurement error associated with stream gage data at extreme low flows. Due to So. CA's heterogeneous landscape, large variations in land use, topography, and precipitation shape flow permanence and flashiness across the region (Table A1). To better discern the effects of these heterogeneities on streamflow, and to more accurately capture time-sensitive environmental flow metrics on a scale relevant to benthic macroinvertebrates, hourly data were chosen over daily. Additionally, high resolution hourly data across So. CA provide an opportunity to complement the previous state-wide classifications (Lane *et al.*, 2017; Lane *et al.*, 2018; Pyne *et al*., 2017), which used daily data, at finer temporal and spatial scales.

Inductive classification was performed to group sites based on similarity of streamflow flashiness (RBI) and permanence (< 1 cfs). To achieve this, a variety of exploratory ordination analyses were conducted to develop an initial understanding of how gages might classify. Weighted classical (metric) multidimensional scaling within the “vegan” package of R (Oksanen *et al.*, 2019) complemented principal component analysis (PCA) and a scree plot from the “stats” package (R Core Team, 2019). Classification was ultimately determined using K-means clustering from the NbClust package in R (Charrad *et al.*, 2014) after assessing the following indices: C-Index, Dunn, McClain, and Silhouette.

2.2.2.2 Deductive Approach

For traditional deductive classification, watershed data describing USGS streamflow gages were retrieved from the USGS’s GAGES-II database (Falcone, 2011) and the U.S. Environmental Protection Agency’s (EPA) NHDPlusV2 database (McKay *et al.*, 2012). Correlation was performed with the “stats” package in R (R Core Team, 2019) to remove highly correlated watershed metrics. Finally, the same exploratory ordination analyses and clustering process as the inductive approach provided results for traditional deductive classification.

2.2.2.3 Combined Inductive and Deductive Approaches

Inductive and deductive methods of stream classification were combined in multiple ways. First, a single K-means clustering analysis was performed using the hydrologic metrics (RBI and < 1 cfs) and the best performing watershed variables from the deductive classification. Next, multinomial logistic regression within the “nnet” package of R (Venables and Ripley, 2002) was used to determine if flow metrics could predict deductively produced clusters, and likewise used to see if landscape metrics could predict inductively produced clusters. Finally, the USGS has categorized streamflow gages containing minimally disturbed watersheds without significant flow alteration as “reference” within the GAGES-II database (Falcone, 2011). Multinomial logistic regression with flow and watershed metrics was again used to predict whether a gage was reference or non-reference.

2.2.3 Hydrologic Model-based Classification

Hydrologic Model-based Classification (HMC) first requires the accurate creation and calibration of rainfall-runoff models across a region, exactly like regionalization for estimating streamflow in ungaged basins. Parsimonious and minimally-calibrated models are important to HMC so that physical relationships between regional watershed variables and highly uncertain model parameters might be established. Rather than using tradition inductive measures of streamflow to assess hydrologic similarity for classification, HMC quantifies the hydrologic similarity between two sites as the reciprocating model accuracy when calibrated parameters from one model are donated to the other and vice versa. Representing hydrologic similarity with model errors produced by a regional range of parameters is a new idea in regionalization that can be used to quantify and reduce parameter uncertainty. Calibrated parameters inherently have greater uncertainty than directly calculated parameters, and this uncertainty is substantially increased in ungaged basins where calibration cannot occur. HMC uses jackknife resampling of complete calibrated parameter sets for all models across the region to generate a model-error matrix of hydrologic similarity spanning the region. The regional error matrix can be interpreted as quantitatively describing parameter uncertainty for the most uncertain parameters across a region. In HMC, the error matrix is used as an inductive basis of hydrologic similarity and combined with a deductive approach to produce a new combined classification that directly incorporates regionalization and reduces parameter uncertainty in models of ungaged basins. Ultimately, classifying models with reciprocally low errors provides a subset of parameters from a calibrated regional catalog with reduced uncertainty. Figure 2.2 provides an overview of the process for HMC.

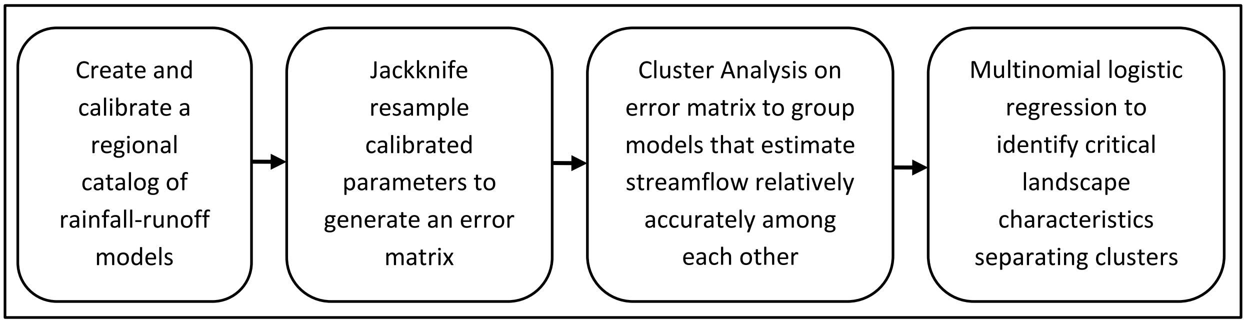

Figure 2.2: Flowchart overviewing novel Hydrologic Model-based Classification (HMC).

2.2.3.1 Hydrologic Models

Hydrological models were created in the US Army Corps of Engineers Hydrologic Engineering Center's Hydrologic Modeling System (HEC-HMS) 4.1 at the 25 gages. Continuous simulations were performed on an hourly time step over WY 2005-2007 to capture a period spanning a wide range of typical hydrologic conditions (WRCC, 2015). Hourly precipitation data were input from the California Irrigation Management Information System (CIMIS), California Data Exchange Center (CDEC), Climate Data Online from the National Oceanic and Atmospheric Administration (NOAA), San Diego County Flood Control District (SDCFCD) and Ventura County Watershed Protection District (VCWPD). CIMIS gages also provided monthly average evapotranspiration data. Independent watershed delineations in ArcMap 10.1 using a 30m digital elevation model from The National Map (USGS, 2019), NHDPlus V2 (McKay *et al*., 2012), and National Land Cover Database (NLCD) (Fry *et al.*, 2011) were verified by USGS StreamStats data (USGS, 2019). Inverse distance was used to weight precipitation gages from each watershed's centroid. Simple canopy (interception and transpiration) and surface (infiltration) parameters were estimated from delineated data. HEC-HMS model parameters associated with the deficit and constant loss element (infiltration) were calculated directly using soil and imperviousness data available from USGS GAGES-II (Falcone, 2011). Similarly, the time of concentration and Clark unit hydrograph storage coefficient used within the Clark unit hydrograph transform element were calculated directly using the Kirpich method (Kirpich, 1940) and standard approaches utilized by the Arizona Department of Transportation (ADOT, 2014). To produce minimally calibrated models, methods were selected to balance simplicity and parameter parsimony with reliable and process-based hydrology. The Kirpich Method, for example, contains only two parameters, which facilitates straightforward calculations in data-scare areas. It is a long-trusted method for estimating time of concentration (USDA NRCS, 2007) that is highly effective across a wide range of conditions in a similar region (Roussel *et al*., 2005).

After directly estimating and calculating parameters associated with precipitation losses and hydrograph transformation, only two linear reservoir baseflow parameters were calibrated for the 25 modeled watersheds. Initial flow values were known using streamflow gage data, and a single linear reservoir was used for each of the two groundwater layers. These two layers were connected in parallel with the both groundwater layers combining to produce a total baseflow (USACE, 2000). As such, only the groundwater storage coefficient for each layer was altered during calibration.

Flashy floods and periods of little precipitation have strongly influenced the evolution of healthy freshwater aquatic ecosystems in So. CA (Gasith and Resh, 1999). In continuing with this study's focus on streamflow flashiness and permanence as ecologically-relevant management metrics, models were calibrated to optimize RBI and < 1 cfs. While the accuracy of a singular measure of overall fit is typically used for hydrologic model calibration (Bardossy, 2007; Beven, 2012), environmental flow studies have shown it is not ideal for modeling ecological flow metrics (Cassin *et al*., 2005; Murphy *et al*., 2013; Parker *et al*., 2019; Vis *et al*., 2015). As a result, calibration accuracy of flashiness and flow permanence were equally considered and combined into one "Ecologically-Focused Combined Calibration" (EFCC), which has been used to calibrate hydrologic models for ecological applications in So. CA (Parker *et al*., 2019). EFCC (Equation 2.4) equally weights the percent error (Equation 2.3) of RBI (Equation 2.1) and < 1 cfs (Equation 2.2).

Equation 2.3: Percent Error.

$$Percent\ Error\ (\%) = \left(\frac{|Gage\ flow\ metric - Modeled\ flow\ metric|}{Gage\ flow\ metric}\right) * 100$$

Equation 2.4: Ecologically-Focused Combined Calibration (EFCC).

$$EFCC\ (\%) = \left[\frac{\left(\frac{|Gage\ RBI - Modeled\ RBI|}{|Gage\ RBI|} \right) * 100 + \left(\frac{|Gage < 1\ cfs - Modeled < 1\ cfs|}{|Gage < 1\ cfs|} \right) * 100}{2} \right]$$

2.2.3.2 Jackknife Resampling Error Matrix

To compute hydrologic similarity among the regional network of minimally calibrated hydrologic models, storage coefficients and initial discharges of both groundwater layers were donated from one model to all 24 remaining models. This was done for every model in the region in a process known as jackknife resampling (Efron, 1982; Friedl and Stampfer, 2014). Model parameters directly calculated or estimated from available landscape data were not jackknifed. Initial baseflow discharges were included in the jackknife analysis and are treated as calibrated parameters because they would be unknown in a PUB analysis. For each individual model's calibrated parameters, jackknife resampling generated 24 time series characterizing streamflow across the region. The accuracy of each simulated hydrograph resulting from jackknifed parameters was assessed by comparing to the 24 observed USGS streamflow gages. The true gage streamflow data do not affect the jackknifing process because they are only used to determine the accuracy of the output flow data resulting from the jackknifed parameters. The accuracy of each jackknifed parameterization was calculated for the entire 25x24 matrix of time series data using the EFCC (Equation 2.4) scaled by minimum and maximum errors, resulting in a normalized 25x24 matrix quantifying the accuracy of each calibrated model when its calibrated parameters were directly input into all other models. Each sites' original calibration error was added to the matrix such that a normalized 25x25 matrix was produced with very small calibration errors spanning the diagonal.

2.2.3.3 Combined Inductive and Deductive Approach

Combining inductive and deductive approaches for Hydrologic Model-based Classification was very similar to the combined approach under traditional classification that implemented multinomial logistic regression. Using the jackknife error matrix of hydrologic similarity, weighted classical (metric) multidimensional scaling, PCA, and a scree plot provided a sense of how sites might cluster. K-means clustering with C-Index, Dunn, McClain, and Silhouette indices was used to split sites into reciprocating low model-error clusters. This inductive approach produced groups of hydrologically similar gages, as measured by a site's ability to accurately model all other sites within its group. A deductive approach was added to HMC by using multinomial logistic regression to determine if watershed variables could predict low-error cluster membership.

2.2.4 Classification Assessment

To better understand the utility of each classification towards estimating flow in ungaged basins, a performance metric dubbed "average cluster error" (ACE) was developed for this study. ACE characterizes the errors produced by donated parameters within a classification method and its classes. Low-error classifications and classes indicate greater certainty in donated calibrated parameters, which inherently contain high uncertainty in models of ungaged basins. Classifications and classes with low ACE values may provide the foundation for accurately modeling ungaged basins with regionalization. ACE was modeled after the cross-validation standard error (CVSE) statistic presented by Wortman (2005) and is displayed in Equation 2.5, wherein *C* is the total number of clusters produced by a specific classification, *c* represents each cluster, *S* is the total number of sites within the given cluster, *s* is each site from the cluster, *Normalized Errors* is taken directly from the jackknife error matrix, and *P* is the total number of sites (25 in this study).

Equation 2.5: Average Cluster Error (ACE).

$$Average\ Cluster\ Error = \frac{\sum_{c=1}^{C}\sum_{s=1}^{S}(Normalized\ Errors_s)}{P}$$

The following example helps explain how Equation 2.5 was used: Say a specific classification divided the 25 sites into 5 equal groups split chronologically (Sites 1-5, 6-10, 11-15, etc.). Total error for the first group would be computed by summing all within cluster errors (when site 1 parameters were applied to Sites 2, 3, 4, and 5; when site 2 parameters were applied to Sites 1, 3, 4, and 5; etc. for site 3, 4, and 5 parameters). This same process would be repeated for the four remaining groups and summed to produce a final total error. The total error would be divided by 25 sites to yield a single metric quantifying the average model error across all sites, exclusive to a specific classification. Following this procedure, ACE values can also be computed for individual clusters unique to one classification, wherein the number of sites assigned to the specific group of interest would take the place of *P* (*P* = 5 when only considering one cluster from the example above), and the $\sum_{c=1}^{C}()$ term would not be used because only one cluster from the classification is considered. Because all sites receiving each model's parameters were treated as ungaged basins during jackknife resampling, the ACE statistic provides insight regarding how well different classifications, or different groups within one classification, might be incorporated into regionalization.

Additionally, the adjusted Rand index (ARI) was computed between each traditional classification technique and Hydrologic Model-based Classification to compare the similarity of any two unique classification. ARI typically ranges from 0 to 1, wherein a value of 0 indicates no similarities between clusters and a value of 1 represents identical clusters; however, negative values can occur if class similarity is less than what would be expected during random clustering (Hubert and Arabie, 1985). Essentially, ARI values near 0 indicate a classification scheme provides unique groups that do not overlap. Specifically, the "clues" package in R (Chang *et al.*, 2010) was implemented to compute an ARI among all suitable classifications.

Between the two measures for assessing classifications in this study, ARI provides an understanding of each classification's ability separate its data, while ACE reflects the ability of a classification, or cluster within a classification, to estimate streamflow in ungaged basins. ARI is a more general metric for insight into data clustering, while ACE is a specific metric focused on cluster performance in ungaged basins. More generally, ARI quantifies between cluster variability while ACE quantifies within cluster variability.

2.3 Results

2.3.1 Traditional Classification

2.3.1.1 Inductive Approach

Classification of hourly flashiness and flow permanence metrics in coastal southern CA resulted in three classes (Figure 2.3). Sites were essentially split according to flow permanence with intermittent streams containing below-average flashiness (Class 1 with 6 sites), perennial streams spanning the full range of flashiness (Class 2 with 10 sites), and ephemeral streams spanning the full range of flashiness (Class 3 with 9 sites). The intermittent class contained the smallest average cluster error with the least within cluster variability (0.2, Figure 2.3), indicating calibrated parameters from models of these streams possessed the least uncertainty. Likewise, the perennial class had the least utility towards ungaged basins because it contained the most within cluster variability (ACE = 0.9, Figure 2.3). When considering all three clusters produced by traditional inductive classification, the ACE was 0.6 (Figure 2.3).

(a)

Class	RBI	< 1 cfs	Avg Cluster Error
1	0.06	0.26	0.2
2	0.10	0.04	0.9
3	0.10	0.75	0.6
All			**0.6**

(b)

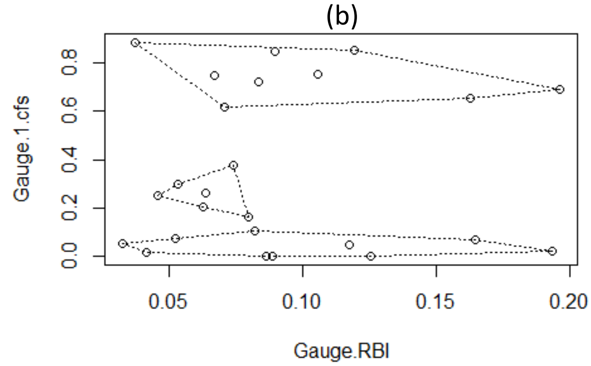

Figure 2.3: Results of inductive approach to traditional classification. Specifically, (a) mean predictor metric values and ACE for the different classes and overall classification; (b) ordination plot illustrating metric values across clusters.

2.3.1.2 Deductive Approach

Classification of watershed characteristics yielded five classes with drainage area and soil content, specifically the percentage of Hydrologic Soil Group C (HGC), providing a parsimonious classification (Figure 2.4). These two watershed variables were log-transformed within the K-means algorithm to address the right skewed nature of drainage area caused by a few large basins. Sites were primarily divided by drainage area, and secondarily by HGC, to generate classes of small basins with low HGC (Class 3 with 3 sites), small basins with high HGC (Class 5 with 7 sites), medium-sized basins with low HGC (Class 1 with 5 sites), medium-sized basins with high HGC (Class 2 with 7 sites), and large basins with high HGC (Class 4 with 3 sites). The large basin with high HGC class contained the smallest ACE (0.2, Figure 2.4), while the medium-sized basin with low HGC provided the largest (0.6, Figure 2.4). An ACE of 0.4 was computed after considering all five clusters produced by traditional deductive classification (Figure 2.4).

(a)

Class	**Drainage Area (km^2)**	**HGC %**	**Avg Cluster Error**
1	143.9	17	0.6
2	206.6	44	0.5
3	24.0	10	0.3
4	1220.6	34	0.2
5	35.4	37	0.5
All			**0.4**

(b)

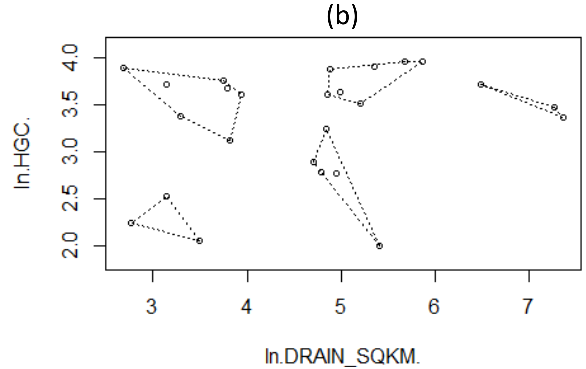

Figure 2.4: Results of deductive approach to traditional classification. Specifically, (a) mean predictor metric values and ACE for the different classes and overall classification; (b) ordination plot illustrating metric values across clusters.

2.3.1.3 Combined Inductive and Deductive Approaches

Results for combining inductive and deductive classification are provided in Appendix A (Figure A2, Figure A3, and Figure A4). Neither an expanded cluster analysis nor predicting inductively and deductively produced clusters with the selected watershed characteristics and flow metrics, respectively, improved classification over the individual inductive and deductive approaches. New multinomial regression models were developed to accurately predict traditional inductive clusters with drainage area, % clay soil, minimum elevation, and annual minimum precipitation, and predict gage reference status with drainage area, % silt soil, baseflow index, and relative humidity.

2.3.2 Hydrologic Model-based Classification

2.3.2.1 Models

Calibration results for the 25 HEC-HMS models at USGS gages are provided in Appendix A (Table A2)Table 3.3. Overall, the flashiness and flow permanence calibration criteria were modeled extremely accurately. Average percent errors of both RBI and < 1 cfs were well under 1%.

2.3.2.2 Combined Inductive and Deductive Approach

Hydrologic Model-based Classification combined inductive and deductive classification to produce a multinomial logistic regression model (deductive classification) that uses landscape variables to predict membership of five hydrologically-similar groups of models (inductive classification) (Figure 2.5). The inductive approach used in HMC does not group sites by the similarity of measured or modeled metrics, as is done traditionally, but instead groups sites to maximize model accuracy when calibrated models' parameters are donated to all other sites within a group. Despite this important distinction, streamflow flashiness and permanence were well distributed across the five hydrologic model-based clusters (Figure 2.5). A multinomial logistic regression model was able to predict low-error class membership with 4% error (24 sites matched correctly) using drainage area, sandy soil content, mean annual precipitation, and mean annual minimum precipitation. The number of sites was distributed less evenly across classes for Hydrologic Model-based Classification than traditional methods, with the first two clusters containing two sites each, the third cluster containing three, the fourth containing five sites, and the final cluster containing over half the sites with 13. As such, it is no surprise that class five contained the largest within cluster variability (ACE = 0.5, Figure 2.5), and is subsequently its worst performing group in ungaged basins. However, no other class within HMC produced an ACE greater than

0.1, which contributes to HMC owning the lowest within cluster variability across all classifications (ACE = 0.3, Figure 2.5).

Stream classes produced by HMC include medium-sized basins with flashiness on both the high (Class 1) and low (Class 4) end. Flashy Class 1 streams receive the least precipitation and are located in southern San Diego County. Non-flashy Class 4 streams comprise the two eastern-most sites. Medium-small basins (Class 3) receive relatively little precipitation and are located near the coast, while large-medium basins (Class 5) receive the most precipitation and are spread throughout the study area. The largest basins (Class 2) are slightly flashier and drier than the large-medium basins (Class 5). These Class 2 streams are concentrated in the northern area of the study area.

Class	Drainage Area (km²)	Sand %	Annual Avg Precip (cm)	Annual Min Precip (cm)	RBI	< 1 cfs	Avg Cluster Error
1	146.4	41	35	1.9	0.16	0.44	0.1
2	463.8	38	51	1.0	0.10	0.33	0.1
3	93.4	33	39	0.6	0.12	0.40	0.0
4	151.9	59	40	1.6	0.05	0.58	0.1
5	222.5	52	55	1.8	0.08	0.28	0.5
All							**0.3**

Logistic Regression Landscape Variable	Definition	Source
DRAIN_SQKM	Total upstream drainage area (km²)	NHDPlus V2 (McKay *et al.*, 2012)
SANDAVE	Percentage of sandy soil (%)	GAGES-II (Falcone, 2011)
PPTAVG_CAT	Mean annual precipitation of NHD catchment (cm)	GAGES-II (Falcone, 2011)
CAT_AnnMinPrecip	Mean annual minimum precipitation of NHD catchment (cm)	GAGES-II (Falcone, 2011)

Figure 2.5: Results of Hydrologic Model-based Classification (HMC).

2.3.3 Adjusted Rand Index (ARI)

The geographical distribution of four unique classifications are displayed in Appendix A (Figure A1), including traditional inductive (flow metrics), traditional deductive (watershed characteristics), a hybrid inductive/deductive (GAGES-II reference sites), and hydrologic model-based as a hybrid inductive/deductive (model accuracy and watershed characteristics). Results of the ARI analysis show no major similarities and large variability between classifications, with the strongest relationship between GAGES-II reference sites and inductive classification (ARI = 0.12, Table 2.1). Inductive and Hydrologic Model-based Classifications were most different with an ARI of -0.04 (Table 2.1).

Table 2.1: Adjusted Rand Index (ARI) among four unique classifications.

	Inductive	**Deductive**	**Reference**	**Hydrologic model-based**
Inductive	-	-0.01	0.12	-0.04
Deductive	-0.01	-	0.004	0.09
Reference	0.12	0.004	-	0.013
Hydrologic model-based	-0.04	0.09	0.013	-

2.4 Discussion

Hydrologic Model-based Classification introduces a new way to think about stream similarity, which can improve the accuracy of hydrologic modeling and environmental flow management in ungaged basins. For hydrologic modeling, HMC can be incorporated into iterative development of a hydrologic foundation and it supplies the foundation for an improved approach to regionalization of ungaged basins. As a management tool, HMC streamlines priority environmental flow metrics in ungaged basins.

2.4.1 Hydrologic Model-based Classification and environmental flow management

Using Hydrologic Model-based Classification to incorporate regionalization for modeling ungaged basins into stream classification provides an opportunity to improve environmental streamflow studies that require ungaged data. ELOHA is an iterative process with significant feedback loops; however, stream classification is recommended to occur second, after developing a hydrologic foundation, and no guidance is provided on how classification might inform the hydrologic foundation or vice versa (Poff *et al.*, 2010). Because the hydrologic foundation generates baseline and current hydrographs at sites with bioassessment data, many of which are ungaged, reciprocally low-error classes produced by HMC could be utilized in a modeling framework to increase the hydrologic foundation's accuracy. Switching the order of the first two steps in ELOHA, and first classifying sites using HMC, could improve streamflow estimation in ungaged basins as a part of the hydrologic foundation. At the very least, developing the hydrologic foundation could be iterative with classification as key characteristics of the sites become better understood, especially if ungaged basins must be modeled.

The primary role of stream classification, as one of the four major steps of ELOHA, is to strengthen and standardize regional flow-ecology relationships so that they may be better implemented for water management (Poff *et al*., 2010); however, it is the one step of ELOHA some studies have determined unnecessary and bypassed (Kendy *et al*., 2012). To this point, large-scale classifications in the Chesapeake Bay watershed (Buchanan *et al*., 2011) and Western US, including a separate classification in California, (Hawkins and Vinson, 2000) did not improve benthic macroinvertebrate explanatory power. While this study has demonstrated how the primary application of stream classification is useful in coastal southern California, it has also introduced HMC to extend classification beyond its traditional role to modeling ungaged basins for developing a hydrologic foundation in any region. It is likely that more accurate hydrologic foundation would create more accurate flow-ecology relationships and stronger environmental flow criteria, and it could also improve the utility of stream classification within ELOHA. This should be evaluated through additional analysis and application.

Modeled streamflow data does not always classify streams the same as gage data for the same sites. Peñas *et al.* (2016) showed daily and monthly gage data clustered better than monthly modeled data in Spain. Similarly, modeled data provided different classes than gaged in North Carolina (Eddy *et al.*, 2017). While model accuracy is always a high priority in hydrologic applications, stream classification is very sensitive to this accuracy, which underscores the importance of accurate models within ELOHA. Poor model accuracy not only directly diminishes the utility of flow-ecology relationships, and subsequent environmental flow recommendations, but it can indirectly hamper management efforts by providing inconsistent stream classes. When ungaged basins are considered in ELOHA, model accuracy must be highly prioritized or else lingering and compounding errors might spoil otherwise legitimate efforts.

From an operational perspective, Hydrologic Model-based Classification is more time-consuming than traditional classifications and might become unwieldy when applied across an expansive geographic region with many sites to classify. This is because not only must hydrologic models be created and calibrated for every classified site, but each model must be analyzed with every other models' calibrated parameters to produce the critical jackknife resampling error matrix. If ungaged basins are to be included, however, some extra time spent building models is recouped as they would have been built anyway under traditional classifications. This study has demonstrated that HMC is feasible for 25 sites spanning a fairly large and highly heterogeneous region in the south coast of California. If a significantly larger region or denser network was the focus of this study, HMC would likely provide even more precise classes and accurate streamflow estimates, but with a substantially greater time investment. Realistically, HMC becomes less feasible at a state-wide scale or for a large network (~50 sites). These issues make HMC most effective when used in concert with large-scale classification methods to enhance classification for relatively small-scale environmental flow development, which might range from basin-level to spanning multiple counties, or with expeditious hydrologic models.

2.4.2 Stream classification for regionalizing ungaged basins

Hydrologic Model-based Classification not only provides new information characterizing regional streams complementary to traditional classifications, but it can also be used to accurately model ungaged basins across heterogenous area through regionalization, as evidenced through the average cluster error metric describing within cluster variability. ACE unpacks important information buried inside the jackknife resampling matrix describing how accurately a set of calibrated parameters can be donated from its original model to all other models in the region, as if the other models were ungaged. Error values from the matrix can be assessed for each model in the region or, when performing stream classification, can be aggregated to quantify ACE for every class within a given classification. Further aggregation can provide an overall measure of ungaged modeling accuracy for an entire classification approach to compare to other classification schemes. A comparison of these overall ACE values shows Hydrologic Model-based Classification containing the least within cluster variability, which provides the most certainty regarding parameters in models of ungaged basins (ACE 0.3; Figure 2.5). HMC was followed by deductive classification with drainage area and HGC (ACE 0.4; Figure 2.4), inductive classification with < 1 cfs and RBI (ACE 0.6; Figure 2.3), and lastly GAGES-II reference status (ACE 1.4; Figure A4).

By providing a method for reducing parameter uncertainty in models of ungaged basins, HMC has demonstrated utility beyond complementary classification. Modeling ungaged basins is fundamental to ELOHA (Poff *et al.*, 2010) and many other hydrology applications, but different approaches vary significantly, contain uncertainty, and do not perform particularly well across a geologically and hydroclimatically diverse area (Arsenault *et al*., 2019; Blöschl *et al.*, 2013). This study provides a foundation for directly incorporating the regional accuracy of a catalog of hydrologic models into a framework for improving ungaged modeling within a heterogeneous region.

This study has shown flow permanence and flashiness were more consistently modeled in ungaged basins containing intermittent streams than ephemeral or perennial streams. Extreme sensitivity to precipitation explains why ephemeral streams did not produce a low ACE, and, while initially, it may be surprising to see baseflow parameters more accurately interchanged between models of intermittent streams than perennial, the effluent nature of perennial streams, especially in a region as rapidly urbanizing as So. CA, inconsistently augments the natural flow regime (Ponce and Lindquist, 1990), and likely prevented accurate modeling in this study. Similarly, flows were modeled with more certainty at GAGES-II reference sites (ACE 0.4; Figure A4) than non-reference (ACE 1.9; Figure A4), wherein flow alteration restricts the ability to transform precipitation into streamflow. Based on the results of this study, intermittent reference streams are likely most accurately regionalized in the south coast.

While no combined classification in coastal southern CA was able to predict class membership of all 25 sites with 100% accuracy, HMC came the closest. This finding underscores the potential for using a measure of model accuracy across a region to define hydrologic similarity within stream classification. Olden *et al.* (2012) split deductive classification into three sub-approaches: “environmental regionalization” to provide a spatial representation of stream similarity, “hydrologic regionalization” using models to estimate flow in ungaged basins, and “environmental classification” for geographically independent classification; however, only one inductive approach, ideal for geographic independence, is described: “streamflow classification”. The new Hydrologic Model-based Classification developed in this study is based on inductive reasoning but is not “streamflow classification”. Instead HMC is a type of “streamflow regionalization” wherein each region is a reciprocally low-error class. Instead of defining geographic areas of assumed flow similarity using watershed characteristics, “streamflow regionalization” directly groups sites based on modeled flow similarity. This new approach essentially hybridizes “hydrologic regionalization” and “streamflow classification”.

Deductive classification produced relatively low uncertainty of model parameters, with all five classes containing ACE values between 0.2 and 0.6 (Figure 2.4). The relatively tight spread coupled with a low overall ACE (0.4; Figure 2.4) implicate deductive classification as a worthy alternative to HMC for regionalization of ungaged basins. These results are consistent with the most common implementation of regionalization wherein models are typically grouped by spatial proximity, physical similarity, or parameter regression (Oudin *et al.*, 2008; Razavi and Coulibaly, 2013; Samuel *et al.*, 2011). This study has shown how a new type of "streamflow regionalization", akin to Hydrologic Model-based Classification, might edge out traditional "hydrologic regionalization" from deductive classification, at estimating streamflow in ungaged basins. "Hydrologic regionalization" and "streamflow regionalization" both implement watershed characteristics to separate sites for high utility in modeling ungaged basins; however, "streamflow regionalization" improves modeling by directly incorporating a quantifiable measure of ungaged model accuracy. This important addition to "streamflow regionalization" directly captures regional model uncertainty and strengthens the science supporting modeling ungaged basins.

2.4.3 Stream classification in coastal southern California

As measured by ARI, traditional inductive classification and reference status classification were the two most similar, but still contained high variability (0.12, Table 2.1). This finding is consistent with how GAGES-II primarily uses flow alteration to classify reference streams (Falcone, 2011), and with how ELOHA recommends classifying by hydrologic similarity to develop flow-ecology relationships (Poff *et al.*, 2010). Furthermore, the reference status classification established a relationship, predominately with drainage area, but also silt content, baseflow index, and relative humidity, which could help water managers identify streams facing potential flow alteration.

The two most different classifications in this study were traditional inductive and hydrologic model-based (ARI -0.04, Table 2.1). Hydrologic Model-based Classification is primarily based on an inductive approach; however, it quantifies hydrologic similarity completely differently than traditional inductive classification. The negative non-random relationship between these classifications is explained as the traditional approach considers gage data similarity and hydrologic model-based considers model data similarity of the same metrics. The differences in these two inductively-based classifications underscore the complexity in modeling streamflow permanence and flashiness in So. CA and suggest great effort must be taken when modeling ungaged basins in the south coast region.

Using ARI, this study has demonstrated how four unique stream classifications can each provide important, complementary information regarding how streams across a region may be grouped for management. While the two inductively-based classifications appear the most useful for separating gaged and ungaged sites, respectively, important relationships and management opportunities can be revealed through a robust regional stream classification using multiple approaches.

2.5 Conclusions

Accurately modeling ungaged basins is often necessary for quantification and management of environmental streamflows (Poff *et al.*, 2010), but it is a difficult undertaking with no consensus approach among the hydrology community, especially in heterogenous regions (Arsenault *et al*., 2019; Blöschl *et al.*, 2013). Furthermore, stream classification is one of the four major steps used to develop environmental flow criteria within ELOHA (Poff *et al.*, 2010), but it is not always used in the framework (Kendy *et al.*, 2012). This study sought to increase the utility of classification within ELOHA while simultaneously strengthening the science supporting modeling and management of ungaged basins in heterogeneous regions. To this end, Hydrologic Model-based Classification was developed to provide: complementary classification information, improved ungaged model accuracy, and new opportunities for stream management. Iterating between the first two steps of ELOHA (hydrologic foundation and classification) within HMC improves both steps and produces stronger environmental flow criteria.

While this study focused on streamflow permanence and flashiness due to their known ecological importance in the study region (Gasith and Resh, 1999; Mazor *et al*., 2018; Parker *et al*., 2019), additional flow metrics corresponding to other element of the flow regime are important in So. CA (Yarnell *et al.*, 2020) and could be incorporated. To develop a better understanding of HMC in general, it could be extended to new regions and compared to the results of this study. This could produce general relationships between different classifications and provide insight into which classification approach might be most appropriate for specific applications and regions. Likewise, a type of nested classification similarly implemented across many regions would help different stakeholders understand how management actions at multiple geographic scales might affect streams and would foster coordinated management relationships. As HMC is expanded to additional regions, a better understanding of the similarity of within-class management plans will be developed. These findings will be highly dependent on the management metrics and regions, but a general sense for management plan transferability within low-error classes will offer a clearer understanding of how Hydrologic Model-based Classification might assist in ungaged stream management without ever modeling the basin.

For coastal southern California, HMC results from this study should be further developed into a full framework for modeling time-series of discharge in new ungaged basin(s) from the heterogenous region. This would foster a better understanding of the modeling complexities within Hydrologic Model-based Classification, and its associated new regionalization framework, and would provide the basis of a hydrologic foundation prioritizing ungaged basins, which is needed to develop robust regional environmental flow criteria in So. CA.

Chapter 3

Estimating streamflows in ungaged basins through regionalization: a new hybrid hydrologic and statistical ensemble modeling framework applied to environmental flow management in a heterogeneous region

Summary

Hydrologic modeling of streamflow regimes in ungaged basins is performed across a wide range of water management applications, including establishing a hydrologic foundation for the development of environmental flow criteria, and yet there is no favored method. Regionalization of rainfall-runoff models is a common approach that estimates unknown parameters for each model of an ungaged basin from a group of regional models. Parameters are often transferred from calibrated model(s) to models of ungaged basins using a measure of geographic proximity, hydroclimatic or landscape similarity, or linear regression. These approaches typically do not incorporate any quantitative data regarding the accuracy of donated parameters, and prove especially difficult in regions with heterogeneous land use, geological settings, and hydroclimatological processes. I address this problem by developing and testing a regionalization framework for estimating streamflow regimes in ungaged basins across heterogeneous landscapes. This new framework, called “Streamflow Regionalization with Hydrologic Model-based Classification” (SR-HMC), implements jackknife resampling and stream classification to directly quantify and reduce regional parameter uncertainty. SR-HMC utilizes a statistical procedure for extrapolating calibrated parameters of rainfall-runoff models from gaged sites to ungaged locations, followed by ensemble modeling to appropriately weight extrapolated parameter sets. The novel framework was tested to predict ecological flow metrics in basins across coastal southern California, and the results were compared to commonly applied nearest neighbor regionalization. Results indicate SR-HMC predicted key flow metrics more accurately (combined median error 16%) than nearest neighbor (combined median errors 20% and 28%). When only nearest neighbor regionalization was considered, models calibrated to specific environmental flow metrics of ecological relevance (extreme low flow and flashiness) outperformed models calibrated using best overall fit. Of the two ecological metrics, low flow indicative of drying was more accurately modeled using the new framework (median 10% error) than streamflow flashiness (median 27% error). While sample size was limited for developing (25 sites) and testing (5 sites) the new framework, coupling jackknife resampling and stream classification technique with ensemble rainfall-runoff modeling was shown to reduce parameter uncertainty and improve the estimation of streamflow in ungaged basins. The time and effort invested into modeling ungaged streamflow with SR-HMC is particularly beneficial when prioritizing the accuracy of predetermined flow metrics across a heterogeneous region.

Keywords: rainfall-runoff modeling, ungaged basins, regionalization, ensemble modeling, jackknife resampling, heterogeneous

3.1 Introduction

Watershed modeling of ungaged basins is used extensively by scientists and engineers to simulate streamflow conditions at locations without measured data for water resources planning and management. Modeling these streamflow predictions in ungaged basins (PUB) is ubiquitous across hydrologic applications (Blöschl *et al.*, 2013), whether for annual events, low flows, large floods, environmental flow management, or anything in between. For environmental flows specifically, PUB is often critical for establishing a hydrologic foundation of currently impaired vs. natural flow conditions for sites with detailed bioassessment data (Poff *et al*. 2010). These crucial bioassessment sites typically occur on wadeable lower order streams, which are inadequately gaged, yet include most of the total stream length in the US (Poff *et al.*, 2006). Unfortunately, streamflow metrics often used for environmental flow studies, such as extremely low flows, can be volatile and have proven difficult to model (Carlisle *et al*., 2010; Blöschl *et al*., 2013; Kennard *et al.*, 2010; Nathan and McMahon, 1992; Razavi and Coulibaly, 2017).

Modeling streamflow in ungaged basins involves a high level of uncertainty and has no consensus method (Arsenault and Brissette, 2014; Bardossy, 2007; Blöschl *et al.*, 2013; Farmer and Vogel, 2013; Hrachowitz *et al.*, 2013; Oudin *et al.*, 2008; Peel and Blöschl, 2011; Razavi and Coulibaly, 2013; Wagener and Montanari, 2011). Some of this uncertainty is unavoidable because watershed models only provide simplifications of reality through mathematical equations that will never completely represent nature. Parameters for these equations can be highly sensitive and contribute heavily to the uncertainty of modeled streamflow (McCuen, 1973). The lack of available discharge data for calibrating parameters in models of ungaged basins results in even more parameter uncertainty than models at gage locations (Blöschl *et al.*, 2013). For best results, regardless of technique, watershed modelers should understand, quantify, and minimize uncertainty (Liu and Gupta, 2007). Understanding and reducing parameter uncertainty is practiced by many modelers through creating parsimonious models, carefully assessing input data, and using multi-objective functions (Wagener *et al.*, 2001), but quantifying and minimizing it remains challenging.

A common method for estimating unknown parameters in models of ungaged basins is regionalization (Tasker, 1982). During typical regionalization, parameters of an existing model calibrated to gage are donated or extrapolated to a model at an ungaged location using one of three primary methods: 1) geographic proximity, wherein parameters from the spatially nearest calibrated model are donated (Kokkonen *et al.*, 2003; Zvolensky *et al.*, 2008); 2) similarity of hydrology, climate, or an ideal landscape characteristic between a calibrated model and an ungaged site determines donated or extrapolated parameters (Li *et al.*, 2009; Zhang and Chiew, 2009); or 3) regional regression to calculate ungaged parameters (Arsenault and Brissette, 2014; Oudin *et al.*, 2008). These simple approaches are less accurate in geographically and hydrologically diverse regions compared to relatively homogeneous regions (Arsenault and Brissette, 2014; Arsenault *et al*., 2019; Blöschl *et al.*, 2013). A region that spans diverse land uses, topography, geology, and/or micro-climatology will likely contain streams with very different flow regimes. Simple regionalization approaches are unlikely to consider such heterogeneity and may inaccurately model streamflow in ungaged basins of such a diverse region.

Within a regionalization framework, ensemble rainfall-runoff modeling, wherein parameters from multiple calibrated regional models are donated to a model of an ungaged basin (McIntyre *et al*., 2005), can be used to reduce parameter uncertainty, as a form of model averaging, by incorporating important parameter features from each donor model within the ensemble (Refsgaard *et al*., 2006). Averaging with ensemble rainfall-runoff modeling typically results in one parameter set derived from the collection of each ensemble model's calibrated parameter set (Burn and Boorman, 1993). Estimating this ungaged parameter set via ensemble modeling can involve two of the three regionalization approaches previously mentioned for a single model: 1) geographic proximity (Yankov *et al*., 2006) or 2) similarity of hydrology, climate, or an ideal landscape characteristic (McIntyre *et al*., 2005). In addition to these simple approaches, ensemble modeling has employed complicated non-linear methods for averaging output discharge, such as Bayesian model averaging (Duan *et al.*, 2007) and artificial neural networks (Tokar and Markus, 2000). As such, no simple method has been developed for quantifying uncertainty across a regional network of models for making ensemble streamflow estimates in ungaged basins, without computationally intensive techniques such as Monte Carlo sampling.

Regionalization of ungaged basins is needed to establish a hydrologic foundation in coastal southern California (So. CA) (Ch. 2; Sengupta *et al.*, 2018; Stein *et al*., 2017) under the Ecological Limits of Flow Alteration (ELOHA) (Poff *et al*., 2010). Yet, the heterogeneity of flow regimes, land uses, and geology across southern CA present a great challenge to simple regionalization schemes, especially when modeling environmental flow metrics. To this end, a new method for classifying streams called “Hydrologic Model-based Classification” (HMC) was developed as a first step towards a more robust regionalization framework (Ch. 2). HMC addresses the challenges of regionalizing flow metrics that are difficult to model in ungaged basins across a highly heterogeneous region by implementing targeting model calibration and directly incorporating regional parameter uncertainty for grouping hydrologically similar streams. While HMC is a useful tool for grouping streams, it needs to be further developed into a full regionalization framework for modeling streamflow in ungaged basins.

The science and methods supporting watershed modeling in ungaged basins have steadily improved but significant challenges remain, especially with estimating environmental streamflow metrics in heterogeneous regions. Accordingly, this research seeks to improve regionalization of streamflow in ungaged basins by quantifying and reducing parameter uncertainty through directly incorporating a measure of regional parameter accuracy and employing ensemble modeling. I hypothesize that these efforts to reduce parameter uncertainty will produce a robust regionalization framework that will outperform nearest neighbor regionalization at modeling environmental streamflow metrics in ungaged basins across a heterogeneous region. Specifically, this study has two objectives:

1) Expand streamflow classification with Hydrologic Model-based Classification (HMC) into a full regionalization framework for modeling time series of discharge in ungaged basins across a heterogeneous region; and
2) Test the framework and compare its performance in modeling environmental flow metrics with common nearest neighbor regionalization approaches.

3.2 Methods

3.2.1 Study Area

Coastal southern California (is a highly heterogeneous region of the United States with some of the most complicated and significant water management issues in the world (State of California, 2019). Land uses span a wide range of urban/suburban, agricultural, and rural coastal and mountainous. Diverse topography, geology, and regional precipitation patterns typical of the semi-arid and Mediterranean climate characterize the area. Due to this substantial regional heterogeneity, streams of all types (perennial, intermittent, and ephemeral) can be found in So. CA. . Recent droughts and wildfires have increased management stress and public awareness of many regional water issues for streams spanning the full range of flow permanence (State of California, 2019). These climactic, hydrogeologic, and water user complexities demand accurate streamflow modeling in ungaged basins and make So. CA as an excellent region to develop and test a new regionalization framework. This research considered the area between the Transverse Mountains, Mexico, the Peninsular Mountains, and Pacific Ocean from coastal regions of counties including San Diego, Riverside, Orange, San Bernardino, Los Angeles, Ventura, and Santa Barbara Counties. The study area is described as the "South Coast" according to the U.S. Geological Survey's (USGS) hydrologic regions of CA (Waananen and Crippen, 1977).

3.2.2 Streamflow Regionalization with Hydrologic Model-based Classification (SR-HMC)

3.2.2.1 Ensemble of rainfall-runoff models

Given the environmental streamflow scope of this research, potential site locations were limited to USGS stream gages near bioassessment sites established by the California Water Boards' Perennial Streams Assessment (PSA) within the Surface Water Ambient Monitoring Program (SWAMP). Locations were screened for dams, diversions, reservoirs, and interbasin transfers that may clearly alter flow conditions between gages and bioassessment sites. Only sub-daily streamflow data were considered for characterizing environmental flow metrics at a time scale relevant to ecological health (Zimmerman *et* *al*., 2010). Gages without sufficient coverage of Water Years (WY) 2005-2007 were eliminated. These three years characterize a representative wet, normal, and dry year, respectively, in So. CA, (WRCC, 2015). WY 2005-2007 provides the recommended reference climatic period for developing environmental flows (Poff *et al*., 2010) and captures an element of heterogeneity in So. CA. Furthermore, land development focused around metro-Los Angeles and San Diego, in concert with the State Water Plan, has led to many artificially augmented and depleted streams. Gages on streams receiving transbasin diversions or downstream of large control structures were omitted. Finally, streamflow gages required high-quality representative meteorological and landscape data to create accurate, yet parsimoniously calibrated models. After a comprehensive search, 30 USGS streamflow gages containing hourly flow data during WY 2005-2007 were selected for this study from the National Water Information System (NWIS) (USGS, 2019b).Of the 30 selected gages, calibrated rainfall-runoff models were developed for 25 to provide a regional ensemble of models, and 5 were withheld for model testing (Figure 3.1; Table 3.1).

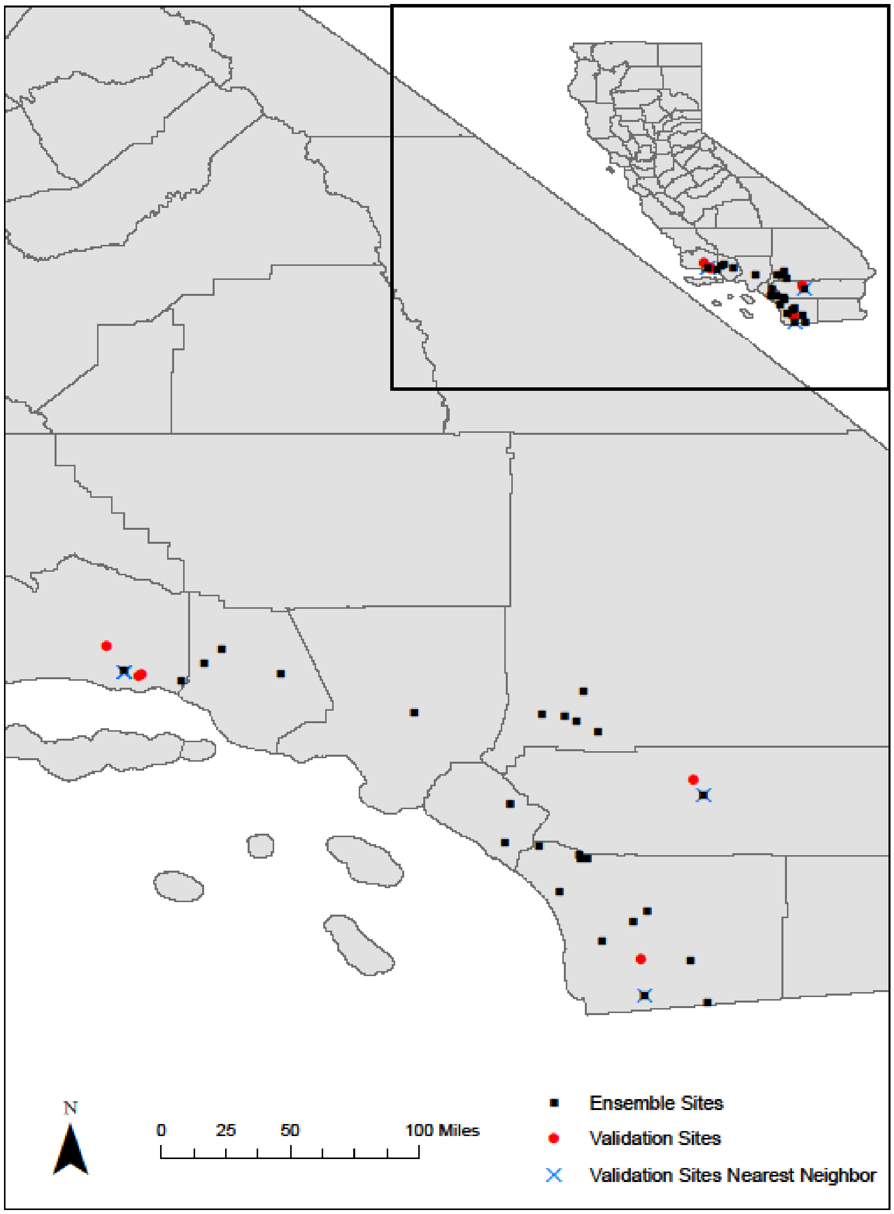

Figure 3.1: Locations of study watershed outlets wherein 25 black squares represent USGS streamflow gage sites used to develop the regional ensemble of rainfall- runoff models; five red circles indicate gage sites used for model testing and validation; and three blue Xs denote the nearest neighbor ensemble model for each validation site (three validation sites have the same nearest neighbor ensemble model).

Table 3.1: Study sites details with ensemble sites indicated by "E" and validation sites by "V". Drainage Area data are from the National Hydrography Dataset Plus Version 2 (NHDPlus V2) (McKay et al., 2012), Impervious data are from StreamStats (USGS, 2019c) and all other landscape metrics are from GAGES-II (Falcone, 2011).

Site Name	USGS Gage	Ensemble or Validation	Drainage Area (km^2)	Imperv-ious %	Sand %	Catchment Annual Avg Precip (cm)	Catchment Annual Min Precip (cm)	Mean Annual Flowrate (cfs)		
								WY 05	WY 06	WY 07
Andreas	10259000	E	23.2	0.0	59.1	32.3	1.0	6.1	2.2	0.9
Arroyo Seco	11098000	E	42.5	0.5	34.1	62.5	1.0	52	8.6	0.9
Arroyo Trabuco	11047300	E	141.4	19.9	32.6	33.8	1.0	69	13	5.1
Campo	11012500	E	222.3	7.0	69.6	41.6	2.0	2.0	0.4	0.1
Carpinteria	11119500	E	45.4	0.1	34.3	44.2	0.0	18	3.7	0.0
Deep Creek	10260500	E	354.0	2.4	63.9	23.4	2.0	171	63	7.8
Devil Canyon	11063680	E	14.7	0.7	53.5	80.3	3.0	2.7	4.9	2.1
East Twin	11058500	E	23.1	0.7	55.6	60.5	2.2	2.7	5.4	1.6
Jamul	11014000	E	182.9	0.5	42.5	36.4	2.0	23	0.1	0.0
Lytle	11062000	E	119.8	0.4	57.5	90.7	2.0	37	32	3.1
Matilija	11114495	E	128.5	0.0	35.6	72.2	0.0	156	37	4.3
Plunge	11055500	E	44.2	1.3	55.4	44.9	2.4	14	7.9	2.1
Poway	11023340	E	110.0	21.8	39.4	33.5	1.0	36	7.3	4.5
Rainbow	11044250	E	27.0	4.3	53.8	45.4	1.0	16	1.4	0.4
San Luis Rey	11042000	E	1433.8	3.1	56.8	31.3	1.0	229	29	9.7
San Mateo	11046300	E	210.2	0.1	47.0	45.5	1.7	90	3.3	0.1
Sandia	11044350	E	51.1	1.3	41.5	45.2	1.0	30	6.1	4.0
San Jose	11120500	E	15.8	0.4	35.2	48.3	0.7	14	3.0	0.2
Santa Margarita Sump	11044300	E	1576.9	4.6	54.8	43.4	1.0	82	18	9.4
Santa Ysabel	11025500	E	290.9	0.1	52.3	43.1	2.0	29	1.7	0.0
Santa Maria	11028500	E	147.7	2.6	51.4	43.9	2.0	16	0.4	0.1
Santiago	11075800	E	32.9	0.1	31.1	49.3	1.0	20	1.8	0.0
Sespe Fillmore	11113000	E	651.0	0.1	36.3	52.4	0.0	515	211	15
Sespe Wheeler Springs	11111500	E	131.9	0.1	36.5	69.0	0.0	87	22	1.4
Sweetwater Descanso	11015000	E	126.0	0.3	61.8	61.9	3.0	21	3.7	1.0
Chino Canyon	10257720	V	12.9	0.2	61.4	40.7	1.0	1.9	0.9	0.0
Los Coches	11022200	V	32.7	9.7	45.1	40.6	1.0	4.7	1.1	0.6
Mission	11119750	V	30.1	4.3	38.7	48.0	0.0	14.0	2.7	0.0

Mission Rocky	11119745	V	30.1	0.7	38.7	48.0	0.0	13.1	2.8	0.2
Santa Cruz	11124500	V	192.3	0.0	34.9	59.4	0.0	60.9	22.4	1.4

For each of the 25 ensemble gages, contributing watersheds were delineated in ArcMap 10.1 with a 30m Digital Elevation Model from the National Map (USGS, 2019a), National Hydrography Dataset Plus Version 2 (NHDPlus V2) (McKay *et al*., 2012), and National Land Cover Database (NLCD) (Fry *et al.*, 2011). These delineations set the foundation for each watershed model and were verified with delineations produced by USGS StreamStats (USGS, 2019c). The U.S. Army Corps of Engineers Hydrologic Engineering Center's Hydrologic Modeling System (HEC-HMS) 4.1 was used to create rainfall-runoff models for each of the 25 ensemble sites. A lumped modeling structure with just one basin per model was used in this study to simplify regionalization through a parsimonious model structure that could still establish a physical understanding of how basin-scale calibrated parameters transfer regionally. Regionalization at the basin-scale is practical for management of entire watersheds and provides a foundation for higher resolution modeling. Hourly precipitation data from the California Irrigation Management Information System (CIMIS), California Data Exchange Center (CDEC), Climate Data Online from the National Oceanic and Atmospheric Administration (NOAA), San Diego County Flood Control District (SDCFCD), and Ventura County Watershed Protection District (VCWPD) were input to continuous simulations at the hourly time scale.

Limiting the complexity and number of parameters in rainfall-runoff models is important to reduce uncertainty (Wagener *et al*., 2001). It is especially important for developing SR-HMC because only the most uncertain parameters, those that cannot be calculated or estimated from available data (i.e. calibrated parameters), are transferred from models of gages sites to models of ungaged sites. Minimizing the number of calibrated parameters creates a parsimonious framework capturing regional uncertainty associated with the calibrated parameters. For this study, only the Single Linear Reservoir Storage Coefficients for each of the two Groundwater layers were calibrated. These storage coefficients are time constants for each layer's single linear reservoir related to the stream's response time (Hydrologic Engineering Center, 2015). Table 3.2 describes the modeling structure for the 25 HEC-HMS ensemble models. Parameters from the Deficit and Constant Loss method were directly calculated using soil and imperviousness data from USGS GAGES-II (Falcone, 2011). Results from these calculations were assessed against delineation data. The Kirpich Method (Kirpich, 1940) was used to calculate Time of Concentration (hr) from delineation data within the Clark Unit Hydrograph Transform method. Finally, a method used by the Arizona Department of Transportation was implemented to compute the Clark Unit Hydrograph Storage Coefficient (ADOT, 2014).

Table 3.2: HEC-HMS model structure for 25 ensemble models.

Model Element	**Element Method**	**Physical Processes or Controls**	**Parameters**	**Calculated Estimated Calibrated or Gage**	**Data**
Precip-itation	Inverse Distance Weighting	Rainfall	Basin Centroid	Calc	StreamStats (USGS, 2019c) National Map (USGS, 2019a) NHDPlus V2 (McKay *et al.*, 2012)
ET	Monthly Average	ET	Rate (in/month) Coefficient	G E	CIMIS
Canopy	Simple	Interception and Transpiration	Initial Storage (%) Max Storage (in) Crop Coefficient Uptake Method	E E E E	StreamStats (USGS, 2019c) National Map (USGS, 2019a) NLCD 2006 (Fry *et al.*, 2011)
Surface	Simple	Infiltration	Initial Storage (%) Max Storage (in)	E E	StreamStats (USGS, 2019c) National Map (USGS, 2019a) NLCD 2006 (Fry *et al.*, 2011)
Loss	Deficit and Constant	Infiltration	Initial Deficit (in) Max Deficit (in) Constant Rate (in/hr) Impervious (%)	E Calc Calc Calc	GAGES-II (Falcone, 2011) StreamStats (USGS, 2019c) National Map (USGS, 2019a) NLCD 2006 (Fry *et al.*, 2011)
Transform	Clark Unit Hydrograph	Topography, Geology, and Land Use	Time of Concentration (hr) Storage Coefficient (hr)	Calc Calc	StreamStats (USGS, 2019c) National Map (USGS, 2019a) NLCD 2006 (Fry *et al.*, 2011) NHDPlus V2 (McKay *et al.*, 2012)
Baseflow	Linear Reservoir	Subsurface	Groundwater 1 Initial (cfs) Groundwater 1 Coefficient Groundwater 1 Reservoirs Groundwater 2 Initial (cfs) Groundwater 2 Coefficient Groundwater 2 Reservoirs	G Calib E G Calib E	NWIS (USGS, 2019b)

While application of this new regionalization framework encompasses general watershed modeling of ungaged basins beyond environmental flows, the ensemble HEC-HMS models used to develop of SR-HMC were calibrated to prioritize ecologically-relevant elements of the flow regime. In the study region, streamflow flashiness and drying are two drivers of local benthic macroinvertebrate assemblage structures (Gasith and Resh, 1999; Mazor *et al*., 2018; Parker *et al*., 2019). Hydrologic alteration of other components of the flow regime affects benthic macroinvertebrates in the region (Yarnell *et al*., 2020). However, calibrating to flashiness and stream drying addresses challenges in modeling ungaged basins because they are difficult to model and are directly influenced by calibrated parameters (Groundwater Coefficients, Table 3.2). As such, a calibration metric (Ecologically-Focused Combined Calibration) (EFCC) (Equation 3.3) (Ch. 2; Parker *et al*., 2019) was implemented which equally weights the accuracy of Richards-Baker Flashiness Index (RBI) (Equation 3.1) (Baker *et al*., 2004) and fraction of discharge time series with flow less than 1 cfs (< 1 cfs) (Equation 3.2) between gage data and models. 1 cfs was chosen as a surrogate threshold indicating a dry stream due to measurement challenges associated with gages as extremely low flows.

Equation 3.1: Richards-Baker Flashiness Index (RBI) (Baker et al., 2004), wherein Q_t *is the discharge at time* t, Q_{t+1} *is the discharge at time step after* t, *and* T *is the final time step at the hourly scale.*

$$RBI = \frac{\sum_{t=1}^{T} |Q_{t+1} - Q_t|}{\sum_{t=1}^{T} Q_t}$$

Equation 3.2: Fraction of time with flow < 1 cfs, wherein $N_{Q<1cfs}$ *is the number of time steps containing streamflow less than 1 cfs and* N *is the total number of time steps containing flow data.*

$$< 1\, cfs = \frac{N_{Q<1cfs}}{N}$$

Equation 3.3: Ecologically-Focused Combined Calibration Criteria (EFCC) as a percent error equally weighting the accuracy of RBI and < 1 cfs between each streamflow gage and HEC-HMS model.

$$EFCC\ (\%) = \left[\frac{\left(\frac{|Gage\ RBI - Modeled\ RBI|}{|Gage\ RBI|}\right) * 100 + \left(\frac{|Gage\ < 1\, cfs - Modeled\ < 1\, cfs|}{|Gage\ < 1\, cfs|}\right) * 100}{2}\right]$$

A common best overall fit metric for hydrologic models, Nash-Sutcliffe Efficiency (NSE) (Nash and Sutcliffe, 1970), was calculated after calibration; however, it was not used as a calibration criterion. NSE lends insight into how well the models developed using EFCC generally represent observed hydrographs, but, because it is biased towards high flows (Jain and Sudheer, 2008; Legates and McCabe, 1999) and does not explicitly consider environmental streamflow elements, NSE was not considered during model calibration.

3.2.2.2 Quantifying regional model accuracy and assigning ungaged sites to ensembles

One novel aspect of this regionalization framework stems from its application of jackknife resampling (Efron, 1982; Friedl and Stampfer, 2014) to quantify regional accuracy and characterize parameter uncertainty. Jackknife resampling within Hydrologic Model-based Classification (Ch. 2) involves donating parameters calibrated to one model to all other models from the region. Using HMC, this regional accuracy was captured as an error matrix, which was applied to cluster reciprocally accurate models into groups for estimating streamflow in ungaged basins with reduced parameter uncertainty. To begin, calibrated Linear Reservoir Groundwater Coefficients and gaged Initial Groundwater Flowrates from one model were donated directly into each of the other 24 ensemble models, as if the 24 were models of ungaged basins. Initial Flowrates from gages were available to calibrate ensemble models but would be unknown for ungaged basins. They must be donated along with calibrated Groundwater Coefficients in this regionalization framework because they are essentially treated as calibrated parameters. For each of the 24 modeled time series, EFCC (Equation 3.3) was calculated to quantify the accuracy of each “ungaged” basin model. This produced 24 values quantifying the accuracy of each of the 24 models when using two parameters that contain high uncertainty in ungaged basins but are known to be accurately calibrated to the 25th model in the region. This process was repeated for all 25 models to create a 25x25 error matrix with each model’s own EFCC calibration error along the diagonal (Table 3.3). Using this error matrix, a series of ordination analyses including principal component analysis (PCA) and a scree plot from the “stats” package of R (R Core Team, 2019), and weighted classical (metric) multidimensional scaling from the “vegan” package (Oksanen *et al.*, 2019), explored how reciprocally low-error sites might cluster. K-means clustering with the NbClust package (Charrad *et al.*, 2015) was used to group reciprocally low-error models after assessing C-Index, Dunn, McClain, and Silhouette Indices. Finally, multinomial logistic regression from the “nnet” package in R (Venables and Ripley, 2002) using openly available landscape data from GAGES-II (Falcone, 2011) and NHDPlusV2 (McKay *et al.*, 2012) identified which physical characteristics best distinguished low-error cluster membership such that an ungaged site could be matched to a cluster.

3.2.2.3 Model averaging

At this stage in SR-HMC, a new ungaged site can be matched to a specific cluster, or ensemble, of models. The matched ensemble provides a group of model parameters that can be exchanged relatively accurately between models from the ensemble; however, there is no sense for how to best donate calibrated model parameters from an ensemble to a new model of an ungaged basin. To investigate this problem, six different model averaging techniques were explored to reduce parameter uncertainty. Half the model averaging techniques involved averaging model inputs (baseflow parameters) and the other half averaged model outputs (time series of discharge). Weighting baseflow parameters directly involved averaging all calibrated parameters from models in a selected cluster and transferring one, ensemble-averaged parameter set. This produced one time series of discharge and is similar to traditional regionalization or estimating parameters by regression. Weighting output discharge required donating intact calibrated parameter sets from each individual member of the cluster. This produced multiple time series of discharge, which were averaged as a form of model (output) averaging, like traditional ensemble modeling. Whether averaging model inputs or outputs, ensembles were each averaged using three approaches: equally weighted, weighted by geographic proximity, and weighted by similarity of the strongest landscape predictor from the multinomial regression equation. Weighting by geographic proximity and landscape predictor involved calculating the Euclidean distance of either the basin centroid (geographic proximity) or strongest landscape predictor from multinomial logistic regression between a model of an ungaged site and each model from a matched cluster. Euclidean distances were used to weight model inputs or outputs such that models from the ensemble with smaller Euclidean distances were given proportionally greater weight.

3.2.3 Testing SR-HMC against common nearest neighbor regionalization approaches

HEC-HMS models for the five USGS streamflow gages withheld for validation were developed in the same manner as the 25 ensemble models (Table 3.2) with two major exceptions: Groundwater Coefficients were not calibrated and Initial Groundwater Flowrates were not included. Details regarding the five validation sites are provided in Table 3.1. The SR-HMC process outlined in Section 2.2 was implemented to generate six time series of discharge at the five validation sites, one time series for each model averaging technique. Environmental streamflow metrics RBI (Equation 3.1), < 1 cfs (Equation 3.2), and EFCC (Equation 3.3), and best overall fit metric NSE, were calculated using modeled and withheld gage data.

Performance of the SR-HMC framework was assessed against two versions of the commonly applied nearest neighbor regionalization approach. First, Linear Reservoir Groundwater Coefficients calibrated to EFCC (Equation 3.3) and gaged Initial Groundwater Flowrates from the spatially nearest ensemble model, as measured by basin centroids, were directly donated to each validation site, respectively. By donating EFCC parameters, this approach prioritizes the accuracy of environmental flow metrics, but does not account for regional heterogeneity. Second, ensemble models geographically closest to validation sites were recalibrated to maximize overall fit according to NSE. These were the only models calibrated to best overall fit in this study. Parameters calibrated to best overall fit were transferred from the geographically nearest ensemble model to each respective model of a validation site. This is a very common PUB approach (Blöschl *et al.*, 2013), but it does not prioritize environmental flow metrics or regional heterogeneity. A general overview of the regionalization framework and assessment is provided in Figure 3.2.

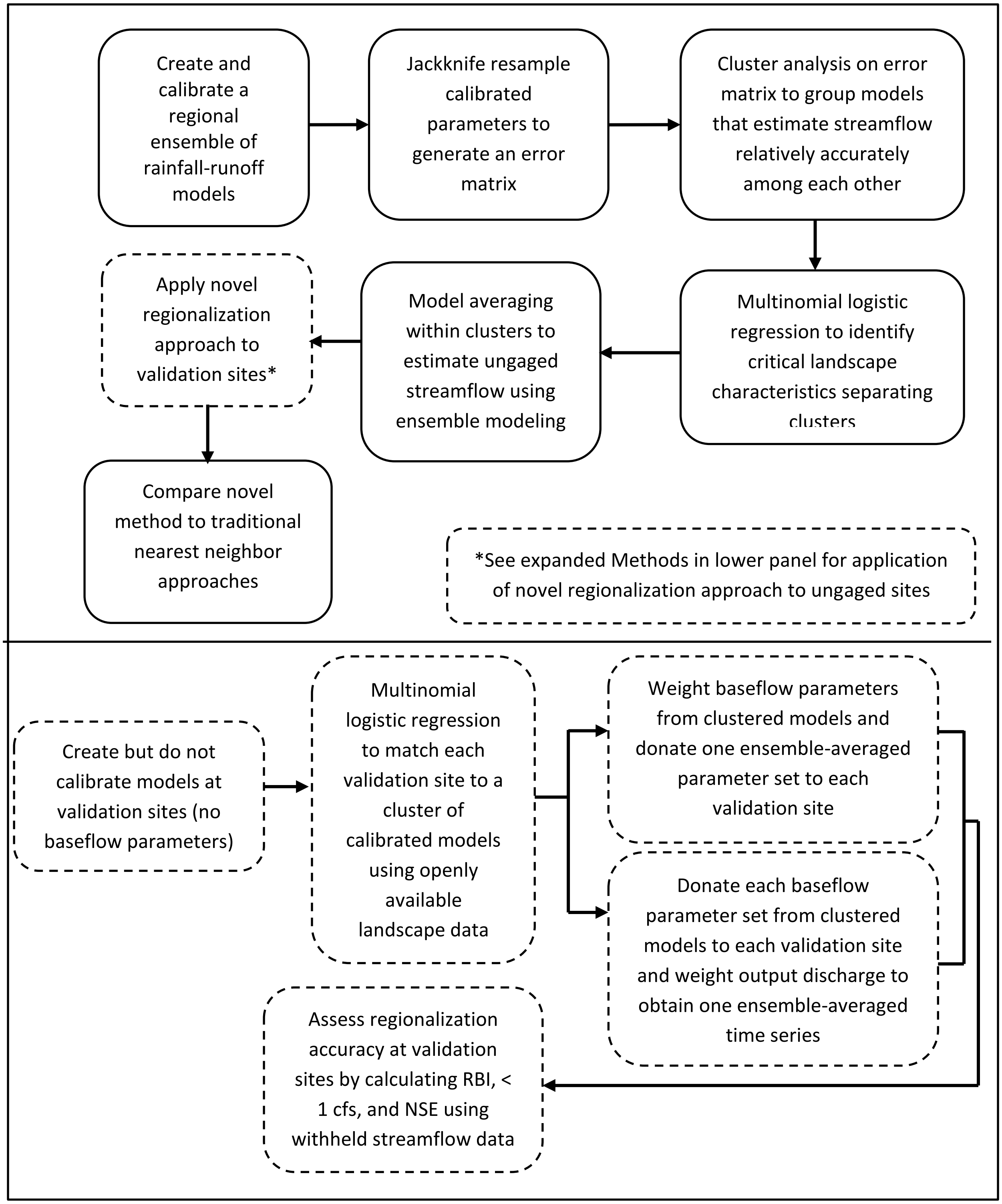

Figure 3.2: Flowchart overviewing the SR-HMC framework (top panel) and application to ungaged sites (lower panel).

3.3 Results

3.3.1 Streamflow Regionalization with Hydrologic Model-based Classification (SR-HMC)

3.3.1.1 Ensemble of rainfall-runoff models

Table 3.3 contains the calibration results for each of the 25 ensemble models. EFCC was very accurately modeled by the regional ensemble, with mean and median errors of 0.3%. Calibrating to environmental flow metrics did not reproduce overall fit or high flow events nearly as accurately, as demonstrated by the NSE values, but despite eschewing any calibration of peak flows in favor of the ecological criteria, nearly 40% of the ecologically-focused models (9/25) produced NSE values greater than 0.3. This relative tradeoff between the accuracy of overall fit and environmental flow metrics was expected. In a preceding environmental streamflow modeling study in So. CA, Sengupta *et al*. (2018) focused on NSE as a calibration criterion with success but was unable to calibrate RBI and < 1 cfs with the accuracy replicated in this study. Furthermore, Parker *et al*., (2019) demonstrated how focusing model calibration on flow metrics that drive ecological response is a more suitable approach for environmental flow applications in So. CA than traditional best overall fit calibration. As such, this study prioritized the accuracy of two important environmental flow metrics (flashiness and stream drying) over all other elements of the flow regime but included a calculation of best overall fit in order to better understand the ramifications of calibration criteria

Table 3.3: Calibration results with model % errors of RBI and < 1 cfs describing model accuracy compared to gaged data. The calibration metric, EFCC (%), was computed using Equation 2.3 Flow records with high flashiness, "H", had a gaged RBI greater than 0.125 during WY2005-2007; sites with low flashiness, "L", had an RBI less than 0.075; sites with average flashiness, "A", had an RBI between 0.075 and 0.125.For flow permanence, ephemeral streams are represented by "E" and had gaged streamflow < 1 cfs more than half the time during WY2005-2007; perennial streams, "P", had flow < 1 cfs less than 10% of time; and intermittent streams, "I", had streamflow < 1 cfs 10%-50% of time. NSE values are given for calibrated models, but overall fit was not considered in the EFCC. Cluster analysis results are included in the final column.

Site Name	Gage RBI	Flashiness	Model % Error RBI	Gage < 1 cfs	Flow Permanence	Model % Error < 1 cfs	EFCC (%)	NSE	Cluster
Andreas	0.05	L	0.0	0.25	I	0.0	0.0	-0.56	4
Arroyo Seco	0.06	L	0.3	0.26	I	0.1	0.2	0.38	2
Arroyo Trabuco	0.16	H	0.1	0.07	P	0.6	0.4	-0.52	3
Campo	0.04	L	0.1	0.89	E	0.1	0.1	< - 2500	4
Carpinteria	0.08	A	0.1	0.72	E	0.0	0.1	0.53	5
Deep Creek	0.12	A	0.9	0.04	P	0.5	0.7	0.28	3
Devil Canyon	0.03	L	0.3	0.05	P	0.6	0.4	-0.52	4
East Twin	0.08	A	0.4	0.16	I	0.2	0.3	-1.7	5
Jamul	0.12	A	0.2	0.85	E	1.1	0.7	-5.9	5
Lytle	0.05	L	0.3	0.30	I	0.1	0.2	-22	2
Matilija	0.05	L	0.3	0.07	P	0.3	0.3	0.50	2
Plunge	0.08	A	0.4	0.10	I	0.0	0.2	-5.6	5
Poway	0.19	H	0.6	0.02	P	1.8	1.2	0.68	1
Rainbow	0.20	H	0.6	0.69	E	0.2	0.4	0.34	5
San Luis Rey	0.04	L	0.2	0.02	P	0.8	0.5	-1.7	1
San Mateo	0.07	L	0.1	0.62	E	0.0	0.1	0.17	5
Sandia	0.09	A	0.9	0.00	P	0.0	0.4	0.74	5
San Jose	0.16	H	0.1	0.66	E	0.0	0.0	0.49	5
Santa Margarita Sump	0.13	H	1.1	0.00	P	0.0	0.6	-29	5
Santa Ysabel	0.11	A	0.5	0.76	E	0.0	0.3	-9.5	5
Santa Maria	0.09	A	0.1	0.85	E	0.0	0.1	-2.2	2
Santiago	0.07	L	0.3	0.75	E	0.0	0.2	0.52	2
Sespe Fillmore	0.09	A	0.3	0.00	P	0.0	0.1	0.31	5
Sespe Wheeler Springs	0.06	L	0.4	0.20	I	0.1	0.2	0.19	5
Sweetwater Descanso	0.07	L	0.6	0.37	I	0.9	0.7	-9.7	5
Mean			**0.4**			**0.3**	**0.3**		
Median			**0.3**			**0.1**	**0.3**		

3.3.1.2 Quantifying regional model accuracy and assigning ungaged sites to ensembles

Hydrologic Model-based Classification (Ch. 2) split the 25 ensemble models into five reciprocally low-error clusters containing between two and 13 ensemble models (Table 3.3; Figure 3.3; Figure B1). HMC also produced a multinomial logistic regression model with four landscape variables predicting cluster membership with an accuracy of 96%, as determined by matching 24/25 models to their correct clusters. These landscape metrics include drainage area, minimum precipitation, average precipitation, and sand soil content, and are further described in Figure B1 and Ch. 2. Based on results of the multinomial logistic regression model, drainage area (DRAIN_SQKM) was the strongest predictor of cluster membership.

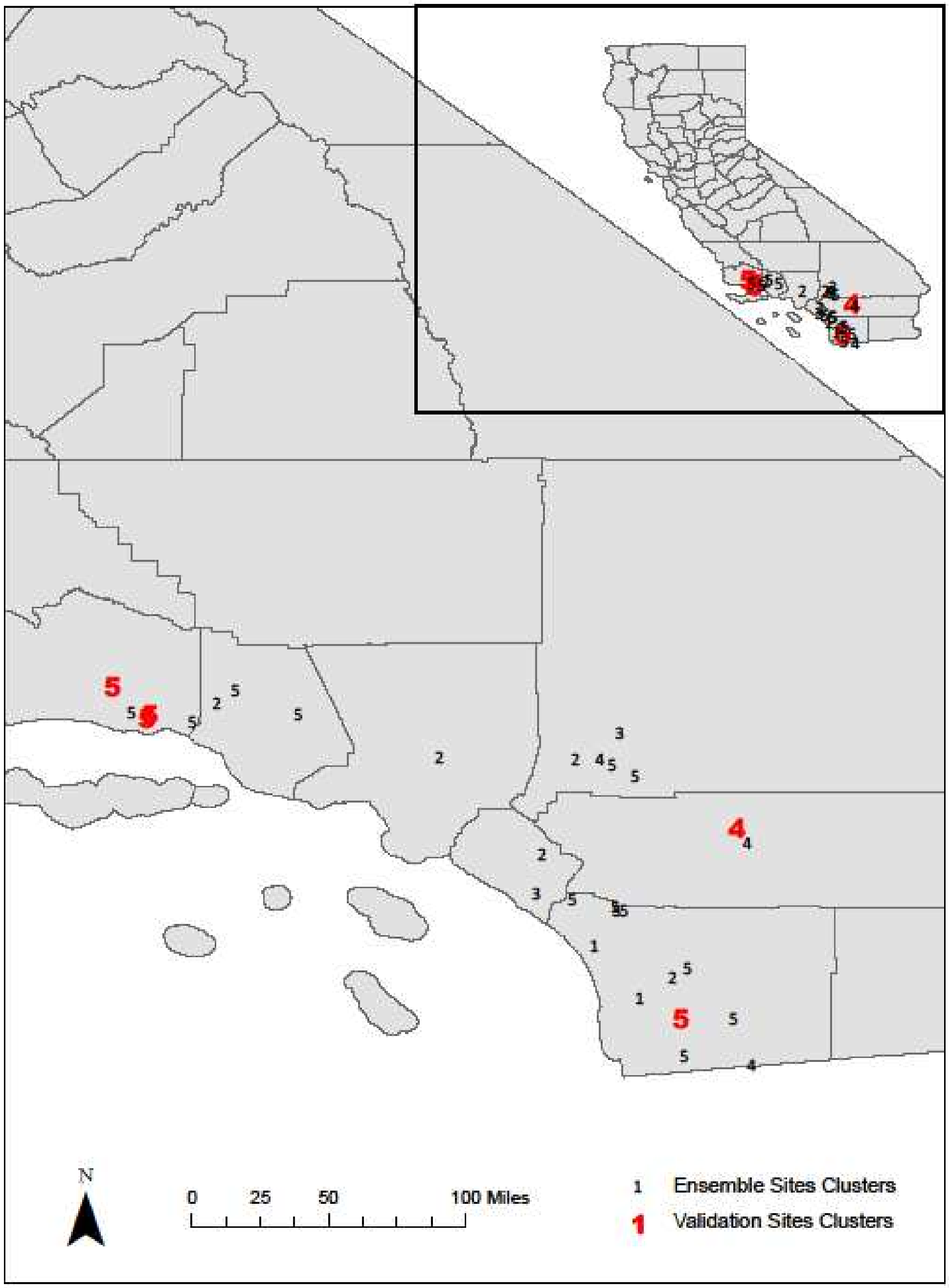

Figure 3.3: Map of HMC clusters with number indicating cluster membership. Smaller, black number represent ensemble sites and larger, red numbers indicate validation sites.

3.3.1.3 Model averaging

The six different averaging techniques tested in the ensemble framework yielded fairly similar, yet appreciably different, results (Table 3.4). In general, each of the three averaging approaches using model output discharge (equal weighting, averaged by best predictor similarity, and averaged by geographic proximity) outperformed their model input averaging counterparts. The ensemble method that most accurately modeled environmental flow metrics across all five validation sites was averaging the output discharges according to geographic proximity (EFCC median error 16%); however, overall fit was poorest for this approach (median NSE 0.27). Conversely, overall fit was maximized when baseflow parameters were scaled by the similarity of DRAIN_SQKM between ensemble sites and ungaged sites (median NSE 0.36), but this improvement in overall fit came at the expense of the least accurate EFCC (median error 30%).

Five of the six weighting schemes performed worst at Santa Cruz (not equal averaging of baseflow parameters). This site uniquely produced some very large errors (> 70%) of < 1 cfs (Equation 3.2). As one of the two components of EFCC, large errors in < 1 cfs at Santa Cruz inflated mean EFCCs for some ensemble techniques. Specifically, one large < 1 cfs error at Santa Cruz (150%) severely skewed the results for averaging output discharge according to geographic proximity. This averaging method provided the smallest overall median EFCC error, but the second largest mean EFCC error (32%). The inflated mean is not truly representative of the accuracy of averaging output discharge by geographic proximity, which actually produced the lowest median error of < 1 cfs (10%) despite the one large error at Santa Cruz. This difference between mean and median highlights the problems using mean to describe small datasets containing an outlier (Ott and Longnecker, 2008), and is why median will be emphasized in this study.

*Table 3.4 Results of the model averaging analysis within SR-HMC. Averaging model output (discharge) is presented above averaging model input (baseflow parameters). See Section 3.2.2.3 (Model averaging) for a description of the six different approaches. *indicates best performing averaging as determined by median EFCC % error (OUTPUT averaged by geographic proximity, 16%).*

		OUTPUT equal averaging				OUTPUT averaged by landscape similarity (DRAIN_SQKM)				OUTPUT averaged by geographic proximity*			
Model	**Cluster**	**RBI (%)**	**< 1 cfs (%)**	**EFCC (%)**	**NSE**	**RBI (%)**	**< 1 cfs (%)**	**EFCC (%)**	**NSE**	**RBI (%)**	**< 1 cfs (%)**	**EFCC (%)**	**NSE**
Chino Canyon	4	44	0.0	22	-13	53	27	40	-9.0	29	20	24	-12
Los Coches	5	19	18	18	-3.1	25	21	23	-3.3	18	5.9	12	-3.5
Mission	5	24	29	26	0.40	12	32	22	0.36	27	4.8	16	0.40
Mission Rocky	5	15	23	19	0.60	4.2	26	15	0.56	20	10	15	0.60
Santa Cruz	5	50	21	36	0.30	47	55	51	0.30	34	150	93	0.27
	Median	24	21	22	0.30	25	27	23	0.30	27	10	16*	0.27
	Mean	30	18	24	-3.1	28	32	30	-2.2	26	39	32	-2.8

		INPUT equal averaging				INPUT averaged by landscape similarity (DRAIN_SQKM)				INPUT averaged by geographic proximity			
Model	**Cluster**	**RBI (%)**	**< 1 cfs (%)**	**EFCC (%)**	**NSE**	**RBI (%)**	**< 1 cfs (%)**	**EFCC (%)**	**NSE**	**RBI (%)**	**< 1 cfs (%)**	**EFCC (%)**	**NSE**
Chino Canyon	4	53	16	34	-8.2	54	27	41	-8.7	36	20	28	-11
Los Coches	5	16	21	19	-2.7	17	26	22	-2.6	20	16	18	-2.8
Mission	5	28	30	29	0.41	24	35	30	0.40	22	14	18	0.40
Mission Rocky	5	21	24	23	0.59	16	30	23	0.58	12	8.5	10	0.60
Santa Cruz	5	35	22	28	0.36	38	97	67	0.36	45	73	59	0.33
	Median	28	22	28	0.36	24	30	30	0.36	22	16	18	0.33
	Mean	31	22	27	-1.9	30	43	36	-2.0	27	26	27	-2.5

3.3.2 Testing SR-HMC against common nearest neighbor regionalization approaches

Nearest neighbor regionalization performed as expected, with the approach calibrated to EFCC simulating environmental flows more accurately than the approach calibrated to NSE (Table 3.5; median EFCC 20% vs. 28%). Likewise, NSE calibrated ensembles regionalized validation sites with a better overall fit than EFCC ensembles (NSE 0.33 vs. 0.09).

In comparing the novel framework to nearest neighbor regionalization, NSE calibrated nearest neighbor regionalization was outperformed in modeling environmental flow metrics by all three model output discharge averaging techniques tested in SR-HMC (Table 3.4; Table 3.5). Nearest neighbor regionalization calibrated to EFCC performed well, but did not model environmental flows more accurately than ensemble averaging by geographic proximity using SR-HMC. The strong performance of EFCC calibrated nearest neighbor regionalization can largely be attributed to its very accurate estimates of the low flow/drying metric, < 1 cfs (Table 3.5).

As with all six model averaging techniques tested in the novel framework, both nearest neighbor approaches suffered from extremely poor performance at Santa Cruz (EFCCs > 120%), which impacted mean errors (Table 3.5). When large errors were given more significance by considering means, NSE calibrated nearest neighbor regionalization universally performed the worst at estimating environmental flow metrics (49% RBI error, 56% < 1 cfs error, and 52% EFCC); however, it produced the best, yet still negative, mean NSE value (NSE -1.3).

Table 3.5: Results of the commonly applied nearest neighbor regionalization approach. The EFCC (Equation 3.3) used to calibrate models in this study was compared to traditional best overall fit calibration using models calibrated to the Nash-Sutcliffe Efficiency (Nash and Sutcliffe, 1970).

		NEAREST NEIGHBOR EFCC environmental flow calibration				**NEAREST NEIGHBOR NSE overall fit calibration**			
Model	**Cluster**	**RBI (%)**	**< 1 cfs (%)**	**EFCC (%)**	**NSE**	**RBI (%)**	**< 1 cfs (%)**	**EFCC (%)**	**NSE**
Chino Canyon	4	28	20	24	-12	60	21	40	-7.0
Los Coches	5	26	13	20	-4.6	29	9.3	19	0.33
Mission	5	0.6	4.5	2.6	0.36	2.5	21	12	0.36
Mission Rocky	5	25	0.8	13	0.58	29	27	28	0.51
Santa Cruz	5	120	120	120	0.09	120	200	160	-0.52
	Median	26	13	20	0.09	29	21	28	0.33
	Mean	40	32	36	-3.1	49	56	52	-1.3

3.4 Discussion

While this overall research is focused on environmental streamflows, the regionalization framework developed in this study can be applied to many ungaged basin modeling scenario. SR-HMC has the flexibility to adapt to any feasible calibration objective function of interest. It is intended for a modeler to choose calibration criteria specific to the models' application(s). As such, calibrating to the EFCC in this study simultaneously applied the new regionalization framework and supported environmental flow development in coastal southern CA. Because SR-HMC was developed with d a calibration criterion designed for environmental flow management in So. CA, assessment of the framework in this study will focus on EFCC and its two components (streamflow flashiness and drying) in a regional context. Developing SR-HMC under these highly specific conditions restricted sample size. As such, this study is effective in demonstrating the SR-HMC framework, but is limited in its ability to reach broad regional conclusions about So. CA.

3.4.1 Regionalization of environmental streamflows in ungaged basins across a heterogeneous area

So. CA is an extremely heterogeneous region of the United States with diverse natural terrain, such as mountains, deserts, and an ocean, and anthropogenic features, including agriculture, expansive urban centers, and complex water management networks, which contribute to a wide range of hydrologic responses and stream types. The challenges associated with estimating streamflow in ungaged basins, particularly through regionalization, in such a diverse region are two-fold: 1) dividing a group of dissimilar sites into relatively homogeneous classes intuitively becomes more difficult as the dissimilarities grow; and 2) accurately predicting streamflow for all three major stream types (ephemeral, intermittent, and perennial) is more difficult than for one, and becomes increasingly difficult for drier streams that are more dependent on rainfall events for flow.

Despite these challenges, SR-HMC performed relatively well at estimating environmental streamflows in ungaged basins across a heterogeneous landscape. The top ensemble averaging method identified for SR-HMC utilized geographical proximity in a type of ensemble nearest neighbor modeling. This suggests two distinguishing elements of SR-HMC are likely responsible for its success: 1) the application of the jackknife error matrix; 2) ensemble modeling of output discharge. Results from this study suggest SR-HMC has the ability to capture and separate much of the regional model and landscape heterogeneity necessary for accurate model predications. In very heterogeneous regions such as So. CA, more data are often required to accurately model environmental streamflow than what is used for nearest neighbor regionalization, likely because even geographically proximate streams have significant hydroclimatic, topographic, or geologic differences. However, even with added data from SR-HMC, geographic proximity was still an important variable for regionalizing ungaged basins. This suggests that in heterogeneous regions where nearby watersheds might differ substantially, they might also contain similarities that provide information to reduce parameter uncertainty and facilitate the accurate transfer of model parameters from gaged to ungaged sites.

Explicitly calibrating rainfall-runoff models with EFCC to optimize environmental streamflow metrics (RBI and < 1 cfs) increased model accuracy of those flow metrics in ungaged basins, compared to best overall fit nearest neighbor regionalization. These results are consistent with findings of other environmental flow studies, such as in Kentucky where NSE was a poor predictor of the entire flow regime needed for management (Murphy *et al.*, 2013) and in Washington (Cassin *et al.*, 2005), the southeastern United States (Vis *et al.*, 2015), and So. CA (Parker *et al*., 2019) where targeted calibration to ecologically-relevant flow metrics have been recommended. These findings support the practice of using a model’s intended application as a guide during calibration to improve model performance (Beven and Binley, 1992;), even when calibrated models are regionalized for ungaged basins. Within the scope of environmental flow management, the utility of models for simulating past, present, and future streamflows is increased when the accuracy of crucial elements of the natural flow regime are identified and prioritized from the beginning of the modeling process.

Estimating low flows in ungaged basins is difficult (Carlisle *et al*., 2010; Razavi and Coulibaly, 2017), especially in semi-arid climates typical of So. CA (Kennard *et al.*, 2010; Nathan and McMahon, 1992). These challenges are largely due to the significant effects of urbanization and the relative volatility of low flows (Blöschl *et al*., 2013). Streamflow flashiness is a newer management metric focused on ecological condition that is also volatile and difficult to regionalize, but the results of this study show that the low flow metric indicative of stream drying (< 1 cfs) was more accurately modeled, on average, than flashiness (RBI) regardless of regionalization approach. In the diverse landscape of So. CA, there are several likely reasons explaining the relative accuracy of extremely low flow over flashiness. Flashiness describes short-term alterations in streamflow, which are controlled very strongly by local, sub-watershed-scale factors and precipitation (Baker *et al.*, 2004). These short-term alterations are much more dependent on land management and localized precipitation within each specific drainage basin than on regional-scale factors. The ensemble of lumped models used to develop SR-HMC provides a simple and straightforward application of regionalization with jackknife resampling of calibrated parameters which captures the broad sub-regional precipitation and land use trends that influence stream drying, but has more difficulty capturing hyper-localized features contributing to flashiness.

Greater uncertainty in identifying localized landscape factors that contribute more to flashiness than stream drying is reflected in the landscape metrics selected for multinomial logistic regression within HMC (Figure B1; Ch. 2). The two precipitation parameters separating reciprocating low error clusters, mean annual precipitation and mean annual minimum precipitation, certainly affect streamflow flashiness, but are more strongly connected to baseflow and stream drying. Both parameters describe overall rainfall volume in a watershed and not local storm intensities, which would likely provide a better understanding of streamflow flashiness. Likewise, the abundance of sandy soil, in part, controls infiltration, which plays a large role in shaping streamflow flashiness and drying, but only soil and infiltration data at relatively coarse spatial scales is easily and widely available (GAGES-II, StreamStats, SSURGO, STATSGO2). At such broad scales, the relative presence of sandy soil may more strongly influence baseflow than flashiness. High-resolution soil and impervious area connectivity data might provide a fuller understanding of how land use and soil type control streamflow flashiness in the study region. The final landscape metric used to distinguish between ensembles, drainage area, scales with streamflow volume (Gotvald *et al.*, 2012), such that streams in smaller watersheds are typically at higher elevations in the mountains of So. CA where hillslopes and channels are steeper, resulting in flashier streams. However, many of these smaller streams are seasonal and dry up more frequently than those in larger watersheds near the coast. All four landscape variables used in the multinomial logistic regression control streamflow flashiness and drying in some capacity; however, they collectively appear to more strongly affect regional baseflow, which is likely the explanation for more accurate estimates of regional < 1 cfs over RBI.

3.4.2 Jackknife resampling with ensemble regionalization of ungaged basins

SR-HMC implements jackknife resampling to create ensembles of reciprocally low-error models that perform well when regionalized. These clusters of hydrologically similar models contain powerful information for reducing parameter uncertainty in models of ungaged basins, which is accessed by multinomial logistic regression. In other regionalization studies, jackknife resampling has been mostly used to cross-validate rainfall-runoff model spatial transferability (Sengupta *et al*., 2018) or to cross-validate statistical (mostly regression) models (Blume *et al*., 2007; Castellarin *et al*., 2004; Castiglioni *et al*., 2009; McCuen, 2005; Sefick *et al*., 2015; Wortman, 2005). As a cross-validation tool, “leave one parameter out” jackknife resampling has been most frequently applied to tune statistical model parameters (Blume *et al*., 2007; Castellarin *et al*., 2004; Castiglioni *et al*., 2009; McCuen, 2005; Wortman, 2005). More recently, withholding all parameters from a site as “leave one site out” jackknife resampling has been developed to assess the overall predictability of intact models (Sefick *et al*., 2015; Sengupta *et al*., 2018). Noori and Kalin (2016) used this concept of “leave one site out” jackknifing to create a regionalization framework that shares some characteristics with the framework developed in this study. Similar to the approach of SR-HMC, Noori and Kalin (2016) connected a rainfall-runoff model

(SWAT) to a statistical model (artificial neural network) trained on "leave one site out" jackknife resampling data.

SR-HMC is made more robust by ensemble streamflow modeling. Results from this study agree with findings of previous studies in southeast Australia (Zhang and Chiew, 2009), across Australia (Viney, 2013), and in Ontario, Canada (Razavi and Coulibaly, 2016), wherein ensembles of multiple models outperformed individual models at estimating streamflow in ungaged basins. In many cases, ensemble modeling likely improves regionalization over an individual model due to the averaging of errors and dissimilarities between sites. While two streams within close proximity of each other are probably similar, they also exhibit differences that will probably reduce model accuracy when parameters are interchanged, especially in a region as hydrologically, topographically, and geologically diverse as So. CA. These differences are likely muted when multiple models are averaged together, generally resulting in more accurately modeled streamflow.

In developing the ensemble averaging scheme for this framework, the three primary principles used to regionalize ungaged basins (geographic proximity to an existing model, similarity of an ideal hydroclimatic or landscape characteristic with an existing model, and regional regression of parameters) were further tested alongside the traditional ensemble approach of model output (discharge) averaging. Results from this study showing model output averaging weighted by geographic proximity outperform regression-based regionalization (as parameter averaging) agree with a large study on over 900 watersheds in France wherein regression-based regionalization performed the worst, and model output averaging using spatial proximity performed the best (Oudin *et al.*, 2008). A study in Quebec, Canada similarly found that model averaging of spatial proximity and physical similarity outperformed regression-based regionalization on over 250 basins (Arsenault and Brissette, 2014). These findings support the ideas of Bardossy (2007) that non-linear relationships between model parameters limit the utility of linear regression for estimating individual model parameters in ungaged basins. Instead, it is recommended that calibrated parameters be considered as one complete set, and not individual parameters, when transferred to an ungaged basin. Within ensemble rainfall-runoff regionalization, averaging the model output (discharge) and not inputs (donated parameters) allows for non-linear interactions of complete parameter sets, which has been shown to improve overall model accuracy.

3.4.3 Limitations to regionalization of ungaged basins

Three ephemeral and one intermittent stream were included among the five validation sites in an effort to directly address the challenges with modeling regional environmental flow metrics in semi-arid and heterogeneous ungaged basins. Only the intermittent stream, Santa Cruz, proved difficult for estimating RBI and < 1 cfs regardless of regionalization approach. Santa Cruz demonstrates some limitations with the regionalization framework, specifically in the heterogeneous landscape of So. CA. Santa Cruz is unique among other sites in this study in that it contains a large and rural watershed (192.3 km^2 and 0% imperviousness, Table 3.2). Low flow indicative of drying was modeled particularly inaccurately for Santa Cruz with < 1 cfs consistently being overestimated. While substantial effort was applied to filter out sites with artificially augmented or depleted flows, it is possible that Santa Cruz can experience a backwater effect from Lake Cachuma such that flow conditions are wetter and less flashy than natural conditions, despite existing in a largely undeveloped watershed. Furthermore, Santa Cruz is located near the edge of the “South Coast” hydrologic region where precipitation and associated hydrological patterns are slightly different than the majority of ensemble sites, which lie squarely in the “South Coast” region. The issues with modeling Santa Cruz indicate how regionalization of ungaged basins is less accurate when pushed to its geographical boundaries, and demonstrate the risk associated with applying nearest neighbor parameter transfer to a site near the boundary of hydrological regions in a large, heterogeneous area. The combination of anthropogenic influence and landscape and climatic heterogeneity at Santa Cruz is poorly captured by the ensemble models, resulting in inaccurate regionalization.

This study focuses on RBI and < 1 cfs in order to address challenges with modeling ungaged basins while also advancing the science supporting regional environmental streamflow development. Flashiness and streamflow drying provide a limited assessment of hydrologic alteration in So. CA. A comprehensive assessment is outside the scope of this study but has been investigated extensively (Mazor *et al*., 2018; Parker *et al*., 2019; Sengupta *et al*., 2018; Stein *et al*., 2017; Yarnell *et al*., 2020; Zimmerman *et al*., 2017).

While the results of this study are encouraging for using SR-HMC to reduce parameter uncertainty in models of ungaged basins across a highly heterogenous region, more studies in different regions will further validate the method. The environmental flow focus of this study coupled with the heterogeneity and urban hydrology of So. CA limited the number of models that could be accurately calibrated to gages across the region. Despite an exhaustive search, major hydrologic alterations and gage proximity to bioassessment sites were issues that eliminated many potential stream gages. As such, this study and Ch. 2 provide a simple introduction and example application of HMC and the SR-HMC framework. Reducing parameter uncertainty with jackknife resampling and grouping of regional calibrated parameters should be applied to a region with a larger sample size to further demonstrate their abilities. A simpler study in a less heterogeneous and developed region without an environmental flow focus could provide more sites for modeling and testing the method. Such a study might provide further clarity on the statistical limits of SR-HMC.

Some highlights of SR-HMC involve its adaptability and emphasis on physical accuracy across a heterogeneous region, balanced by an ease of understanding and limited computational intensity with lumped rainfall-runoff models. However, if a modeler desired a more spatially detailed regionalization with SR-HMC, semi-distributed rainfall-runoff models with subbasins, or even fully-distributed models on spatially discrete grids, could be used. These different model structures would complicate low-error clusters, which would no longer be simple ensembles of intact lumped parameters, but these complicated clusters would still provide a more targeted and far less computationally intensive method for quantifying and reducing highly complicated regional parameter uncertainty than using a Bayesian (Beven and Binley, 1992; Thiemann *et al*. 2001) or Metropolis Algorithm (Kuczera and Parent, 1998; Vrugt *et al*., 2008) approach. While the high dimensional Monte Carlo resampling algorithms employed by these computationally intensive methods for reducing uncertainty can be constrained by estimated probabilities of parameter accuracy, they generally search for model parameters from each parameter's full range of possible values. Instead of optimizing parameters from their full range of values for one model, SR-HMC optimizes parameters from a subset of parameters calibrated to other models in the region and can do this for any model. In this study, manual calibration of just two co-dependent parameters (Groundwater 1 and 2 Coefficients) provided highly accurate ensemble models (Table 3.3). If this framework was applied to a scenario with more uncertainty surrounding the calibrated parameters of ensemble models, a Monte Carlo approach for reducing parameter uncertainty at each ensemble site could precede SR-HMC.

3.5 Conclusions

Modeling streamflow in ungaged basins is difficult, and frequently inaccurate (Blöschl *et al.*, 2013), yet is often crucial for a wide range of hydrologic and engineering applications. This study sought to improve regional streamflow estimation in ungaged basins across a heterogeneous region by developing a framework incorporating a measure of regional model accuracy. To this end, SR-HMC combines HEC-HMS rainfall-runoff modeling with a statistical procedure of jackknife resampling, cluster analysis, multinomial logistic regression, and ensemble model averaging, to increase the accuracy of modeled streamflow in five ungaged basins over commonly applied nearest neighbor regionalization. SR-HMC was developed to prioritize the accuracy of environmental flow metrics, which are difficult to model, across the heterogeneous landscape of coastal southern CA. The initial development and testing of SR-HMC has been applied on a small scale to environmental streamflow management; however, the practice of incorporating model error into regional flow estimations can be extended to any stream management application. While it certainly requires more time to develop this new regionalization than to build a single nearby calibrated model, it may be worth the effort to increase model accuracy in such heterogeneous regions as So. CA. Within the context of rainfall-runoff modeling, jackknife resampling offers a powerful tool for generating data describing how accurately parameters from one model can be transferred to another model. When jackknife resampling is incorporated into Hydrologic Model-based Classification and combined with ensemble modeling, models can be grouped across a region by how accurately their parameters can be interchanged, and then extrapolated as a group to an ungaged basin under the full Streamflow Regionalization with Hydrologic Model-based Classification framework.

This regionalization framework builds on stream classification efforts in So. CA (Ch. 2) to introduce new ideas to the watershed management and modeling communities, while also exploring the first two steps of the ELOHA framework (Poff *et al.*, 2010). Future work in CA should continue within ELOHA by applying SR-HMC to estimate historical streamflow conditions, and their deviations from current conditions, before ultimately creating and assessing flow-ecology relationships. Extending the new framework to estimating flow-ecology relationships, and comparing those relationships to those produced by other common techniques for estimating streamflow in ungaged basins, would be beneficial in advancing the science supporting environmental flow management, and for developing environmental flow criteria in coastal southern CA.

More generally, future work with SR-HMC could dive deeper into its sensitivity to calibration and error matrix assessment criteria. This study demonstrated the framework using models calibrated to environmental flow metrics to predict the same metrics, but applying SR-HMC with less specialized calibration criteria to predict a wider range of flow metrics would provide valuable information. Application of SR-HMC outside the scope of environmental streamflows would likely reduce some restrictions for selecting sites, which would provide more sites for creating regional ensembles and fully test the method. A variety of hydroclimatic regions could be analyzed and assessed in this manner to establish an even broader understanding of the framework. As it builds credibility through regional testing, SR-HMC, and especially jackknife resampling, could be automated within existing rainfall-runoff models. This would naturally facilitate expanding the regionalization framework to semi- or fully-distributed models for more detailed applications.

Chapter 4

Quantifying hydrologic alteration and flow-ecology relationships in ungaged basins for environmental streamflow management in heterogeneous regions

Summary

Managing water resources to incorporate environmental streamflows involves strategic manipulation of flow regimes to improve the ecological condition of streams. To facilitate environmental flow management, regional alterations of natural flow regimes are analyzed alongside changes in ecological benchmarks. This analysis typically begins by pairing bioassessment sites with flow alteration data. Unfortunately, most bioassessment sites are established on small, wadeable streams that lack representative gage data. In the absence of gage data, hydrologic models of ungaged basins are typically used to estimate flow alteration, often through regionalization of models at gaged locations. Modeling streamflow predictions in ungaged basins (PUB) is difficult and has substantial uncertainty. It is especially challenging to estimate environmental flow metrics at ungaged basins in complex regions with highly heterogeneous, land use, topography, geology, and/or hydroclimatology, where geographically nearby basins may have very different hydrologic signatures. Inaccurate models affect the integrity of flow-ecology relationships, yet simple methods for PUB and quantifying flow alteration are typically used to develop flow-ecology relationships. I address these problems by testing the sensitivity and relative accuracy of flow alteration and flow-ecology relationships across a group of five case study sites. Flow alteration and flow-ecology relationships produced by my novel regionalization framework for modeling environmental streamflows in ungaged basins across a heterogenous region, called

“Streamflow Regionalization with Hydrologic Model-based Classification” (SR-HMC) were compared to two versions of nearest neighbor regionalization, one calibrated to overall fit and the other to ecologically-relevant flow metrics. For the two environmental flow metrics analyzed (streamflow flashiness and drying/permanence) across five sites, flow permanence predictions were less sensitive to choice of regionalization method, flow alteration, metric, and precipitation. The two flow metrics differ appreciably in their sensitivities, and judicious application of them as management endpoints requires future research. Understanding the regional sensitivities of flow alteration and flow-ecology relationships to hydrologic modeling variables can help produce robust environment streamflow criteria.

Keywords: environmental streamflow management, ungaged basins, hydrologic modeling, regionalization, ELOHA, flow-ecology relationships

4.1 Introduction

Environmental streamflow management can play an important role in the sustainable use of freshwater resources by reestablishing and maintaining key elements of natural flow regimes (Poff *et al*., 1997). The Ecological Limits of Hydrologic Alteration (ELOHA) (Poff *et al*., 2010) framework for developing regional environmental flow criteria is used to establish relationships between streamflow alteration and ecological conditions that inform environmental streamflow standards. To generate these flow-ecology relationships, a hydrologic foundation of currently impaired vs. pre-altered or “baseline” flow conditions is created for locations containing bioassessment data. Data from stream gages can be used for the hydrologic foundation; however, hydrologic models are frequently needed to estimate both impaired and baseline flow conditions (Poff *et al*., 2010).

When developing environmental flow criteria, hydrologic modeling most often occurs in ungaged basins where flow conditions at bioassessment sites are not collected. Regionalization of ungaged basins from a network of process-based models at gaged locations (Tasker, 1982) is a common approach used to make streamflow predications in ungaged basins (PUB) across hydrologic applications (Blöschl *et al.*, 2013), including for environmental flows (Ch. 2; Ch. 3; Buchanen *et al*. 2013; Kennen *et al*., 2008; Sengupta *et al*., 2018). Regionalization of ungaged basins is an expeditious approach when developing environmental streamflow criteria with ELOHA because it shares the regional approach (Ch. 2). Regionalization typically involves the transfer of calibrated model parameters from models at gage locations to models at ungaged locations, where calibration data do not exist. This parameter transfer provides insight regarding dominant physical hydrologic processes, and most often occurs between the geographically nearest models (Arsenault and Brissette, 2014; Blöschl *et al.*, 2013; Kokkonen *et al.*, 2003; Zvolensky *et al.*, 2008), but this nearest neighbor approach can be challenging in regions containing streams with diverse flow regimes resulting from heterogenous land uses, topography, geology, and/or micro-climatology (Arsenault and Brissette, 2014; Arsenault *et al*., 2019; Blöschl *et al.*, 2013).

To regionalize impaired streamflow in ungaged basins, models should first be developed at gage locations and calibrated to flow records from a time period that spans natural climactic conditions and includes alterations to the natural flow regime (Poff *et al*., 2010), often a period from the recent past. These parameters calibrated to impaired flow conditions are transferred from model(s) at gaged locations to ungaged locations through regionalization. Previous environmental flow studies recommend targeted calibration that specifically addresses physical elements of the flow regime important to ecological health, instead of best overall fit model calibration (Cassin *et al*., 2005; Murphy *et al*., 2013; Parker *et al*., 2019; Vis *et al*., 2015). For baseline streamflow, hydrologic models are needed not only when nearby stream gage data does not exist for a bioassessment site, but also if gage data exists but does not go back in time far enough to capture historical conditions representative of the natural flow regime. For this task, model parameters that control physical hydrologic processes to represent current streamflows conditions can be adjusted to characterize baseline streamflow

(Buchanen *et al*., 2013; Sengupta *et al*., 2018; State of Colorado, 2016). While this study focuses on regionalization with process-based hydrologic models that provide physical connections between model parameters and watershed characteristics, other studies considering very large geographic areas are less concerned with physical hydrologic processes and have used statistical models to simulate baseline streamflow (Zimmerman *et al.,* 2017).

After estimating baseline and altered streamflows, available software can quantify flow alteration of environmental flow metrics at each bioassessment site (Bledsoe *et al.*, 2007; Henriksen *et al.*, 2006; Richter *et al.*, 1996), typically by computing the alteration of individual flow metrics as a percent deviation from baseline (Poff *et al*., 2010), or % Flow Alteration (Table 4.1). Another approach used in environmental streamflow studies for computing flow alteration measures an altered flow metric as a fraction of its baseline (Carlisle *et al.*, 2010a; Carlisle *et al.*, 2010b; Zimmerman *et al.*, 2017) and can be thought of as an Alteration Ratio (Table 4.1). The most basic approach for quantifying flow alteration involves simply subtracting baseline conditions from altered without any normalization, as Simple Alteration (Table 4.1). These measures of flow alteration are crucial to environmental flow management, yet they are often applied without judicious consideration of their differences.

Table 4.1: Methods for computing flow alteration from altered (or current) and baseline (or historic) streamflow data.

Flow Alteration Metric	**Equation**	**References**
% Flow Alteration	$\frac{(\text{Altered} - \text{Baseline})}{\text{Baseline}} * 100\%$ **Or** $\frac{(\text{Current} - \text{Historic})}{\text{Historic}} * 100\%$	Buchanan *et al.*, 2013 Kennen *et al.*, 2013 McManamay *et al.*, 2013 Poff *et al.*, 2010 Poff and Zimmerman 2010
Alteration Ratio	$\frac{\text{Altered}}{\text{Baseline}}$ **Or** $\frac{\text{Current}}{\text{Historical}}$	Carlisle *et al.*, 2010a Carlisle *et al.*, 2010b Zimmerman *et al.*, 2017

Alteration	(Altered − Baseline) **Or** (Current − Historical)	Mazor *et al.*, 2017 Stein *et al.*, 2017b Sengupta *et al.*, 2018

For any flow alteration metric, hydrologic modeling of ungaged bioassessment sites is often a critical component of developing flow-ecology relationships to inform environmental streamflow criteria; yet at best, these models are only simulated representations of natural systems useful for guiding water management decisions (Beven, 1989). No matter the level of detail included in hydrologic models, they will, by definition, never completely accurately simulate reality. Subsequently, any differences between models, methods, or changes to a model's structure, such as calibration criteria, parameter estimation, input data, etc., will affect estimated streamflows, along with any derived metrics and ensuing estimates of flow alteration (Ch. 3; Duan *et al.*, 2006; Franchini and Pacciani, 1991; Murphy *et al.*, 2013; Parker *et al.*, 2019; Reed *et al.*, 2004; Vis *et al.*, 2015. While regionalization of ungaged basins is a well-established approach that aligns with ELOHA, its traditional application is not well suited for highly heterogenous regions. The effects of modeling decisions on flow alteration metrics and flow-ecology relationships, especially in ungaged basins modeled with regionalization, warrants further investigation.

Coastal southern California (So. CA) is one such heterogeneous region where conflicts between human development and stream health has induced stream alteration (Stein *et al*., 2012) and facilitated ELOHA for restoring regional ecological conditions (Ch. 2; Ch. 3; Mazor *et al.*, 2018; Parker *et al.*, 2019; Pyne *et al.*, 2017; Sengupta *et al.*, 2018; Stein *et al.*, 2017a; Stein *et al.*, 2017b). Two elements of the natural flow regime in So. CA have been identified as important influencers of ecological condition, as determined by local benthic macroinvertebrate assemblages: streamflow flashiness and drying (Gasith and Resh, 1999; Mazor *et al.*, 2018; Parker *et al*., 2019). As such, these two flow metrics have been used in ELOHA efforts across the region and will be the focus of this study; however, they are not the only ecologically-relevant flow metrics. More recent research in So. CA used a functional flows approach to first identify elements of the annual hydrograph important for ecological processes, and subsequently quantify relevant flow metrics (Yarnell *et al*., 2020). New work in So. CA has led to the development of a novel approach for regionalizing environmental streamflow in ungaged basins across a heterogenous region, called “Streamflow Regionalization with Hydrologic Model-based Classification” (SR-HMC) (Ch. 3). SR-HMC combines ensemble rainfall-runoff modeling with a unique stream classification scheme, Hydrologic Model-based Classification (HMC) (Ch. 2), to prioritize model accuracy across a region. Within ELOHA, this recent work generated impaired environmental streamflow conditions, setting the stage to advance ELOHA by computing baseline flow conditions, estimating deviations between current and baseline flows, and developing relationships between flow alteration and changes in ecologic condition.

Modeling streamflow predictions in ungaged basins often plays a crucial role in environmental streamflow management, but important modeling decisions are typically made without a clear understanding of their impacts on flow alteration and flow-ecology relationships. These issues are particularly pronounced when modeling environmental flow metrics in heterogenous regions. I address these challenges in this study with three objectives:

1) Compute alteration of streamflow with SR-HMC and compare its accuracy and consistency to common nearest neighbor regionalization approaches;
2) Analyze the sensitivity of flow-ecology relationships in ungaged basins to three modeling choices: method for regionalizing ungaged baseline streamflow, flow alteration metric, and modeling time period; and
3) Provide management recommendations for robust flow-ecology relationships in ungaged basins.

4.2 Methods

4.2.1 Study Area

This study builds off previous environmental streamflow work in So. CA using many of the same paired bioassessment and USGS streamflow gage sites (Ch. 2; Ch. 3; Mazor *et al.*, 2018; Parker *et al.*, 2019; Sengupta *et al.*, 2018; Stein *et al.*, 2017a). Specifically, the five validation gages used to test SR-HMC (Ch. 3) are further analyzed for flow alteration and flow-ecology relationships in this study. Figure 4.1 displays a map of the sites and Table 4.2 provides watershed, streamflow, and bioassessment data. Pairing bioassessment sites with minimally impacted stream gages limited the number of sites available. As such, the five validation gages analyzed in this study and Ch. 3 do not fully characterize regional flow alteration and flow-ecology relationships, but instead represent a series of case study sites from which general conclusions can be drawn about modeling flow alteration and flow-ecology relationships in ungaged basins across a heterogenous region.

Geographically, the greater study area ranges approximately 230 miles from the Mexico border to the Transverse Mountains north of Santa Barbara. From the Pacific Ocean, the area lies roughly within a 60-mile coastal band constrained to the east by the Peninsular Mountain Ranges. Heavy development has occurred in this region of the United States with, the 2010 census recording over 21,000,000 people (U.S. Census Bureau, 2012) living in regional counties (San Diego, Riverside, Orange, San Bernardino, Los Angeles, Ventura, and Santa Barbara). While urban land use is extensive in So. CA, significant agricultural and rural land, particularly at higher elevations, also exist. Over the decades, persistent population growth and land use change, combined with sub-regional topography, geology, and precipitation distributions, have created tremendous water management issues (State of California, 2019). Due to these external pressures, streams of all types (ephemeral, intermittent, and perennial) have deviated substantially from their natural flow and sediment regimes, which has affected channel form and stability (Stein *et al*., 2012) and led to a decline in overall ecological health (Mazor *et al*., 2017; Stein *et al*., 2017a).

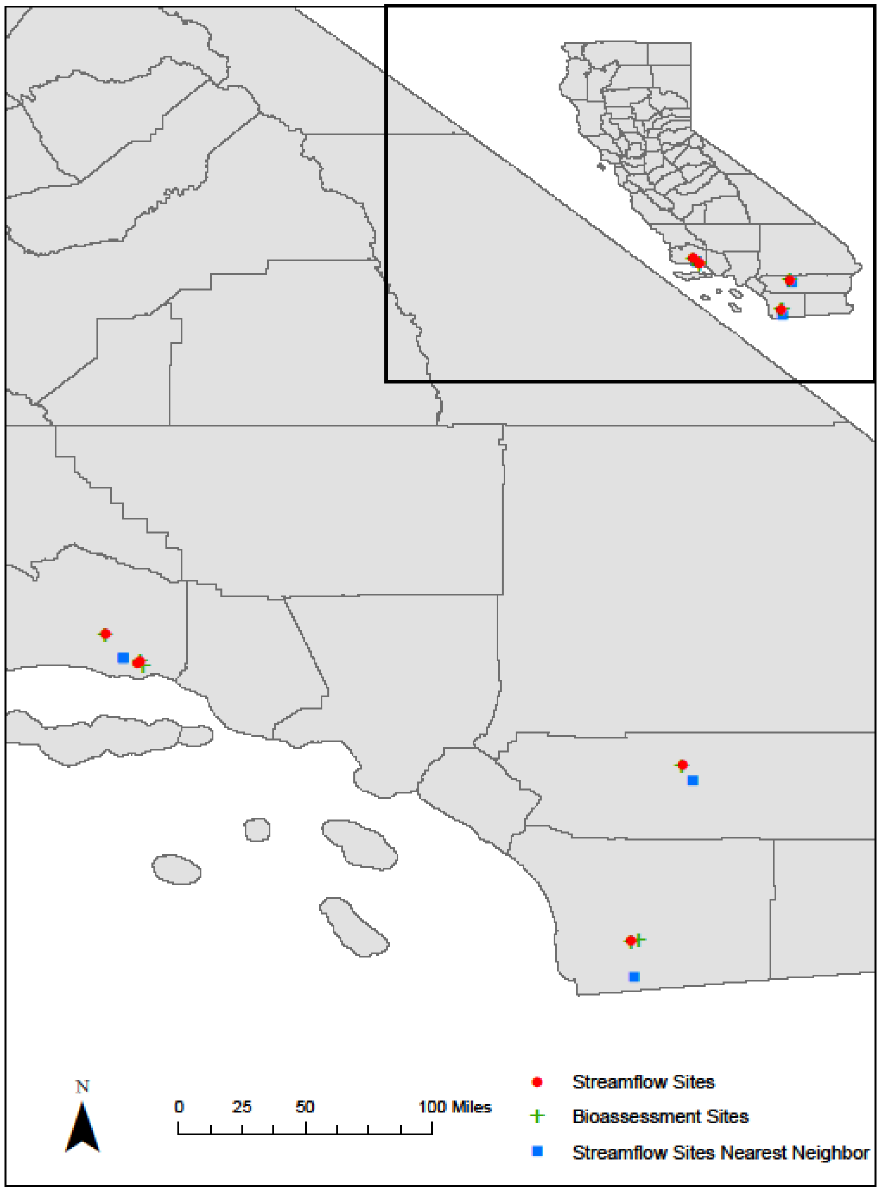

Figure 4.1: Study region with five streamflow sites indicated by red circles (same as Validation Sites in Ch. 3, their associated bioassessment sites labeled with green plus signs, and each streamflow site's nearest neighbor HEC-HMS models labeled with blue square.

Table 4.2: Study sites details. Drainage Area data are from the National Hydrography Dataset Plus Version 2 (NHDPlus V2) (McKay et al., 2012). Flow data with high flashiness, "H", had a gaged Richards-Baker Flashiness Index (RBI) (Baker et al., 2004) greater than 0.125 during WY2005-2007; sites with low flashiness, "L", had an RBI less than 0.075; sites with average flashiness, "A", had an RBI between 0.075 and 0.125. For flow permanence, ephemeral streams are represented by "E" and had gaged streamflow < 1 cfs more than half the time during WY2005-2007; and intermittent streams, "I", had streamflow < 1 cfs 10%-50% of time. Bioassessment data are from the California Water Boards' Perennial Streams Assessment (PSA) within the Surface Water Ambient Monitoring Program (SWAMP). Values of CSCI less than one indicate degraded ecological condition (Mazor et al., 2016).

Site Name	USGS Gage	Oldest Gage Record	Drainage Area (km^2)	WY 2005-2007 Flashiness	WY 2005-2007 Flow Permanence	Bioassessment Site(s)	CSCI
Chino Canyon	10257720	WY 1989-1991	12.9	L	E	719TRMDSS	1.09 1.02 0.91
Los Coches	11022200	WY 1989-1991	32.7	H	E	907LCCHW8	0.94
						907S11430	0.85
Mission	11119750	WY 1989-1991	30.1	H	E	315MISxxx	0.77 0.61 0.52 0.52
Mission Rocky	11119745	WY 1999-2001	30.1	A	E	315MIU	1.17 1.09 1.06
Santa Cruz	11124500	WY 1993-1995	192.3	L	I	314SCCNSY	0.78

4.2.2 Altered streamflow at gage sites

When a stream gage lies close enough to a bioassessment site(s) such that their flows are highly similar, half of flow alteration can be quantified by simply pulling the record for a representative climactic period during altered, likely current, flow conditions. In part, the five gages chosen for this study were selected due to their proximity to bioassessment sites and accurate instantaneous flow record for a typical wet, average, and dry year in coastal southern CA (Water Years, WY, 2005-2007) (WRCC, 2015). For this study, hourly time series from WY 2005-2007 were used as altered streamflow at gage sites. Due to the significance of streamflow flashiness and permanence on determining the ecological health of coastal southern CA streams (Gasith and Resh, 1999; Mazor *et al.*, 2018; Parker *et al.*, 2019), and challenges associated with modeling them in ungaged basins (Blöschl *et al.*, 2013; Carlisle *et al.*, 2010a; Nathan and McMahon, 1992; Razavi and Coulibaly, 2017), flow metrics characterizing these two elements were computed. Specifically, flashiness was quantified by the Richards-Baker Flashiness Index (RBI) (Equation 4.1) (Baker *et al*., 2004) and drying was quantified by the fraction of discharge less than 1 cfs (< 1 cfs) (Equation 4.2). A threshold of 1 cfs was chosen to indicate a dry stream due to the difficulties gaging extremely low flows.

Equation 4.1: Richards-Baker Flashiness Index (RBI) (Baker et al., 2004), wherein Q_t *is the discharge at time* t, Q_{t+1} *is the discharge at time step after* t, *and* T *is the final time step at the hourly scale.*

$$RBI = \frac{\sum_{t=1}^{T} |Q_{t+1} - Q_t|}{\sum_{t=1}^{T} Q_t}$$

Equation 4.2: Fraction of time with flow < 1 cfs, wherein $N_{Q<1cfs}$ *is the number of time steps containing streamflow less than 1 cfs and* N *is the total number of time steps containing flow data.*

$$< 1\, cfs = \frac{N_{Q<1cfs}}{N}$$

4.2.3 Baseline streamflow at gage sites

Two estimates of historical baseline streamflow were generated at gage sites. First, the oldest three WY on record were selected for the five sites (Table 4.2), from which RBI and < 1 cfs were computed. For the second estimate, rainfall-runoff models were created for current flow conditions (WY 2005-2007) and relevant parameters were adjusted to reflect baseline conditions. For this task, the U.S. Army Corps of Engineers Hydrologic Engineering Center's Hydrologic Modeling System (HEC-HMS) 4.1 was used. Models were created and calibrated to optimize a metric combining RBI and < 1 cfs, the Ecologically-Focused Combined Calibration Criteria (EFCC) (Ch. 2; Ch. 3; Parker *et al*., 2019), as displayed in Equation 4.3:

Equation 4.3: Ecologically-Focused Combined Calibration Criteria (EFCC), which equally weights the accuracy of RBI and < 1 cfs between each streamflow gage and HEC-HMS, as a percent error.

$$EFCC\ (\%) = \left[\frac{\left(\frac{|Gage\ RBI - Modeled\ RBI|}{|Gage\ RBI|} \right) * 100 + \left(\frac{|Gage\ < 1\, cfs - Modeled\ < 1\, cfs|}{|Gage\ < 1\, cfs|} \right) * 100}{2} \right]$$

Methods of model creation and calibration for WY 2005-2007 followed the approach outlined for "ensemble models" in Ch. 3, which should be referenced for additional detail.

With individual models accurately calibrated to EFCC for the five sites, applicable land use parameters were adjusted to reflect historic conditions. Given the substantial regional landscape heterogeneity in So. CA, and the judgement required to reduce land use parameters to represent predeveloped hydrology, relatively simple but reliable approaches were taken to model historic conditions. Specifically, imperviousness was set to 0% to represent predeveloped conditions without paved surfaces. Additional loss parameters were not changed. Paving land was assumed to capture the most significant changes to post-development hydrology, while non-paved land use was assumed similar between pre- and post-development. Additionally, the longest flow path was increased to augment Time of Concentration (TOC) (Kirpich, 1940) and the Clark Unit Hydrograph Storage Coefficient (ADOT, 2014). Adjusting the length of the longest flow path reflects the extra distance and slower speed traveled by runoff from the hydrologically most distant location without the effect of impervious surfaces, which significantly increase runoff compared to natural terrain. The amount by which the longest flow path was increased was iterated on to increase TOC by 15%, which is consistent with the literature (Chin, 2006; Iowa SUDAS, 2013; Masch, 1984; Schueler, 2000, and USDA, 2010). RBI and < 1 cfs were calculated from these models of baseline conditions at gage sites.

4.2.4 Altered streamflow in ungaged basins

In a preceding study, SR-HMC for regionalizing ungaged basin was developed and tested against "nearest neighbor" regionalization to estimate environmental streamflow (RBI and < 1 cfs) for current conditions at the five sites over the same representative WY 2005-2007 (Ch. 3). With regionalization of ungaged basins, parameters from nearby calibrated models are transferred to uncalibrated models (Tasker, 1982). For nearest neighbor regionalization, parameters from the geographically closest calibrated model are transferred (Blöschl *et al.*, 2013; Kokkonen *et al.*, 2003; Zvolensky *et al.*, 2008). While typically applied to models of ungaged basins, this technique can be applied to sites with gage data for the assessment of methods for modeling ungaged basins. In doing this, gage data is withheld until after modeling streamflow. In this study, the hydrologic modeling results of Ch. 3 will be used for the ungaged altered streamflow component of stream alteration. Nearest neighbor models were calibrated to two different criteria, EFCC (Equation 4.3), and best overall fit, as measured by the Nash-Sutcliffe Efficiency (NSE) (Nash and Sutcliffe, 1970) using USGS streamflow data. Only the best performing ensemble averaging approach of SR-HMC tested in Ch. 3 was considered (output discharge averaged by geographic proximity). This approach combines measures of geographic proximity, regional parameter uncertainty, and physical watershed similarity to average multiple output time series of modeled ungaged basins. Each time series of discharge corresponds to calibrated parameters donated from a model at a gage site within the region. See Ch. 3 for more detail on SR-HMC.

4.2.5 Baseline streamflow in ungaged basins

Modeling ungaged basins is a very common problem across hydrologic applications, most often applied to solve current issues with current land uses. This study extends new developments in PUB to modeling historical conditions in ungaged basins, which is not such a common task. While still withholding gage data to consider the five sites as ungaged, this study modeled historical streamflow by decreasing imperviousness and increasing the length of the longest flow path by 20% within the SR-HMC and nearest neighbor regionalization approaches of Ch. 3. This was the same process used for models at gages (2.3 Baseline streamflow at gage sites). RBI and < 1 cfs were computed from the final averaged time series of discharge representative of historical baseline conditions for each of the five sites.

4.2.6 Formulating flow alteration

Five approaches for assessing flow alteration were analyzed among two different scenarios: gaged and ungaged (Table 4.3). For the gaged scenario, WY 2005-2007 gage data were paired with the two estimates of baseline streamflow (oldest gage record and historically adjusted rainfall-runoff models). These two gaged measures of flow alteration were computed to provide “reference” values for ultimately assessing the accuracy of ungaged flow alteration and flow-ecology relationships. For the ungaged scenario, baseline and altered conditions were computed by SR-HMC, EFCC calibrated nearest neighbor, and NSE calibrated nearest neighbor regionalizations. For each of the five approaches to quantify flow alteration (Table 4.3), Alteration Ratio and Simple Alteration were estimated in accordance with the equations in Table 4.1. Additionally, the % Flow Alteration concept was applied as a fraction three ways, normalized by baseline, altered, and an average between altered and baseline flows, to produce five metrics of flow alteration per approach in Table 4.3. In total, 10 estimates of flow alteration were made for the gaged scenario, as five formulations of flow alteration (three % Flow Alterations, Alteration Ratio, and Alteration) x two baseline estimates (Oldest 3 WY on record and historical land use applied to WY 2005-2007 models). When sites were treated as ungaged basins, 15 values of flow alteration were produced, as the five formulations of flow alteration x three regionalization approaches (SR-HMC, nearest neighbor EFCC, nearest neighbor NSE). The 25 estimates of gaged and ungaged flow alteration were computed for the five sites (Table 4.2) to produce 125 unique quantities of flow alteration for analysis. These quantities of flow alteration were produced for two environmental streamflow metrics, RBI and < 1 cfs, resulting in 250 estimates of flow alteration for creating flow-ecology relationships.

Table 4.3: Methods for estimating altered and baseline streamflow conditions at the study sites under gaged and ungaged scenarios. [1]Three consecutive Water Years typifying wet, average, and dry conditions in So. CA (WRCC, 2015). [2]Environmentall-Focused Calibration Criteria (Equation 4.3). [3]from Ch. 3.

Gaged

Current Altered	Historical Baseline
1) WY 2005-2007[1]	1) Oldest 3 WY on record
2) Models of WY 2005-2007 calibrated to EFCC[2]	2) Historical land use applied to WY 2005-2007 models

Ungaged

Current Altered	Historical Baseline
1) Nearest neighbor regionalization calibrated to EFCC[3]	1) Historical land use applied to nearest neighbor regionalization calibrated to EFCC
2) Nearest neighbor regionalization calibrated to NSE[3]	2) Historical land use applied to nearest neighbor regionalization calibrated to NSE
3) SR-HMC	3) Historical land use applied to SR-HMC

2.7 Quantifying flow-ecology relationships

For each of the five sites, bioassessment data were obtained from the California Water Boards' Perennial Streams Assessment (PSA) within the Surface Water Ambient Monitoring Program (SWAMP). The locations of bioassessment sites can be found in Figure 4.1 with more information provided in Table 4.2. A multimetric bioassessment index known as the California Stream Condition Index (CSCI) was correlated with each of the 10 gaged and 15 ungaged estimates of flow alteration for RBI and < 1 cfs. Multimetric indices are ideal measures of ecologic condition that combine multiple data from biological surveys, including both taxa- and trait-based metrics, into one unitless value that assesses the ecological health of a site against regionally estimated baseline conditions (Barbour *et al*., 1995). CSCI was developed to capture deviations from benchmark reference ecological condition for all streams regionalized across the state (Mazor *et al.*, 2016; Rehn *et al.*, 2015). Specifically, CSCI provides a normalized index to assess the degradation of ecological condition by combining a measure of taxonomic completeness, the ratio of observed-to-expected benthic macroinvertebrate taxa (O/E), with a measure of ecological structure using a predictive multimetric index (pMMI). In this study, when multiple bioassessment sites were paired to one of the five study gages, or multiple estimates of CSCI were made over time for the same bioassessment site, both mean and median values of CSCI were analyzed.

4.3 Results

4.3.1 Altered and baseline streamflow

Table C1 includes HEC-HMS land use parameters used to calibrate models to currently altered streamflow conditions, alongside modified parameters for historical baseline flows. Increasing the length of the longest flow path by 20% for pre-urban conditions increased Time of Concentration by 15%, in accordance with the literature (Chin, 2006; Iowa SUDAS, 2013; Masch, 1984; Schueler, 2000, and USDA, 2010).

*Table 4.4: Altered and baseline RBI and < 1 cfs generated by gaged and ungaged approaches. *Modeled in Ch. 3.*

	Gaged Altered				Gaged Baseline			
	WY 2005-007 from record		Models calibrated to EFCC for WY 2005-2007		Oldest three WY from record		Historical land use in EFCC WY 2005-2007 models	
	RBI	< 1 cfs	RBI	< 1 cfs	RBI	< 1 cfs	RBI	< 1 cfs
Chino Canyon	0.07	74.7	0.07	74.7	0.04	95.5	0.06	74.8
Los Coches	0.22	79.1	0.23	78.8	0.15	89.7	0.07	80.5
Mission	0.17	77.7	0.17	77.7	0.27	97.0	0.13	80.0
Mission Rocky	0.11	75.4	0.12	75.4	0.12	78.9	0.11	76.4
Santa Cruz	0.06	32.2	0.06	32.2	0.06	31.6	0.06	32.2

	Ungaged Altered*						Ungaged Baseline					
	Nearest neighbor regionalization calibrated to EFCC for WY 2005-2007		Nearest neighbor regionalization calibrated to NSE for WY 2005-2007		SR-HMC for WY 2005-2007		Historical land use in nearest neighbor regionalization calibrated to EFCC for WY 2005-2007		Historical land use in nearest neighbor regionalization calibrated to NSE for WY 2005-2007		Historical land use in SR-HMC for WY 2005-2007	
	RBI	< 1 cfs	RBI	< 1 cfs	RBI	< 1 cfs	RBI	< 1 cfs	RBI	< 1 cfs	RBI	< 1 cfs
Chino Canyon	0.05	59.9	0.03	59.3	0.05	59.9	0.04	60.7	0.02	60.1	0.04	61.9
Los Coches	0.28	89.6	0.29	86.4	0.26	83.8	0.15	92.2	0.01	88.9	0.14	68.7
Mission	0.17	74.2	0.17	93.8	0.12	81.4	0.14	76.4	0.14	97.0	0.10	69.4
Mission Rocky	0.14	76.0	0.15	95.8	0.09	83.2	0.14	77.0	0.14	97.0	0.10	70.9
Santa Cruz	0.14	71.6	0.14	96.6	0.09	81.2	0.14	71.6	0.14	96.6	0.10	43.7

4.3.1.1 Gaged sites: altered and baseline streamflow

For the hydrologically representative period of WY 2005-2007, altered streamflow flashiness and drying from the gage record were calibrated with high accuracy in hydrologic models (Table 4.4). Only at Los Coches did modeled RBI and < 1 cfs deviate from gage values (gage RBI and < 1 cfs = 0.22 and 79.1; model RBI and < 1 cfs = 0.23 and 78.8; Table 4.4).

Gaged baseline environmental flow metrics for the oldest three WY on record and the EFCC calibrated model parameterized with historical land use are dryer and less flashy than WY 2005-2007 conditions (Table 4.4). Between the two estimates of gaged baseline conditions, the oldest three WY on record is much dryer for three of the five sites and is slightly wetter for one of the sites, compared to modeling with historical land use. For all five sites, baseline streamflow estimated by oldest three WY on record closely followed local climate trends that were not representative of the long term climate, as indicated in Figure C1. The three sites with WY 1989-1991 as the oldest on record were extremely dry, as was WY 1989-1991. Similarly, Mission Rocky was moderately dryer than average during WY 1999-2001 and Santa Cruz was wetter during WY 1993-1995. This sensitivity of environmental streamflow to precipitation is expected and is why an emphasis was placed on a representative hydrologic period. As such, historically parameterized EFCC calibrated models for the representative climatic period (WY 2005-2007) will be considered the most accurate “reference” gaged estimate of historical baseline environmental streamflow in this study. Accordingly, all errors associated with ungaged historical flow, flow alteration, and flow-ecology relationship will be in reference to the results produced by the historical EFCC model.

Table 4.5: Percent errors for modeling ungaged historical environmental flow conditions with different regionalization approaches. Errors are in relation to historically parameterized EFCC calibrated models for the representative climatic period (WY 2005-2007). +Errors of the oldest three WY from the gage record also provided.

	Ungaged baseline streamflow conditions											
	Nearest neighbor regionalization (EFCC)			Nearest neighbor regionalization (NSE)			SR-HMC			Oldest gage record+		
	RBI % error	< 1 cfs % error	EFCC	RBI % error	< 1 cfs % error	EFCC	RBI % error	< 1 cfs % error	EFCC	RBI % error	< 1 cfs % error	EFCC
Chino Canyon	30.8	18.9	24.8	63.2	19.7	41.4	34.5	17.2	25.9	42.5	27.7	35.1
Los Coches	113.6	14.6	64.1	86.8	10.4	48.6	98.5	14.7	56.6	124.0	11.5	67.7
Mission	4.8	4.5	4.6	8.6	21.2	14.9	20.4	13.3	16.8	108.5	21.3	64.9
Mission Rocky	26.0	0.8	13.4	30.5	26.9	28.7	5.2	7.3	6.2	11.2	3.2	7.2
Santa Cruz	121.5	122.2	121.8	123.4	200.1	161.7	61.8	35.8	48.8	9.4	2.0	5.7
Median	30.8	14.6	24.8	63.2	21.2	41.4	34.5	14.7	25.9	42.5	11.5	35.1
Mean	59.3	32.2	45.8	62.5	55.7	59.1	44.1	17.7	30.9	59.1	13.1	36.1

4.3.1.2 Ungaged basins: altered and baseline streamflow

Model regionalization results for altered and baseline environmental streamflows in ungaged basins are displayed in Table 4.4, below the gaged results. Results for altered ungaged conditions are from Ch. 3.

A more detailed comparison of the three ungaged regionalized baseline estimates of RBI and < 1 cfs are included in Table 4.5, where it can be seen that nearest neighbor regionalization calibrated to EFCC slightly outperformed SR-HMC, as measured by median EFCC error (24.8% vs. 25.9%). However, nearest neighbor regionalization calibrated to EFCC performed extremely poorly at Santa Cruz, resulting in a much higher mean EFCC error (45.8% vs. 30.9%). Regardless, these two regionalization approaches explicitly calibrated to the environmental flow metrics selected for management were more accurate at modeling environmental flows than best overall fit regionalization. Overall, historical stream drying was consistently modeled more accurately than historical flashiness in ungaged basins (Table 4.5).

4.3.2 Flow alteration formulation

Table C2 displays gaged and ungaged values of all 250 flow alteration metrics for streamflow flashiness and drying. For most metrics, ungaged modeling with SR-HMC and nearest neighbor regionalization calibrated to EFCC were similarly accurate to modeling historical land use with gage data. Nearest neighbor regionalization calibrated to best overall fit is clearly the worst performing regionalization approach for computing environmental flow alteration. Independent of gaged or ungaged approach, most flow alteration metrics produced in this study have values less than one and tend to be extremely small. Only values of Alteration Ratio are consistently near one, making them easier to interpret.

4.3.3 Flow-ecology relationships

Flow-ecology relationships between CSCI and gaged and ungaged environmental flow alteration metrics are provided in Table 4.6. Flow alteration data in Table C2 were used to produce the correlation relationships. Because only five sites were analyzed, major conclusions about regional flow-ecology relationships are limited and interpretation of the results will focus more on the direction of correlation coefficients than magnitude. Regardless of flow alteration metric, flow-ecology relationships resultant from modeling historical land use under the gaged scenario were starkly different than those produced from the oldest gage record.

For streamflow flashiness, modeling historical land use showed a decrease in ecological condition associated with an increase in altered RBI. Essentially, as streamflow conditions become flashier following urbanization, ecological condition declines. For the ungaged scenario, these results were only reproduced by SR-HMC. When the oldest gage record was considered for baseline streamflow, strong correlation in the opposite direction was produced. Using nearest neighbor regionalization, correlations in this same opposite direction, albeit of smaller magnitude, were produced by all but one flashiness alteration metric (not Alteration computed by EFCC nearest neighbor; Table 4.6).

Results of streamflow drying in ungaged basins were consistent across regionalization approaches and were of similar magnitude and the same direction as modeling historical land use. This positive correlation between < 1 cfs alteration metrics and CSCI can be interpreted as a decrease in ecological condition associated with a decrease in stream drying after urbanization. Interestingly, no correlation was found between stream drying and CSCI using the oldest gage record for baseline conditions of flow alteration.

*Table 4.6: Spearman correlation, ρ, flow-ecology relationships between CSCI and different RBI and < 1 cfs flow alteration metrics for gaged and ungaged modeling approaches. Red bold coefficients are the opposite direction as the relationships generated for the most accurate "reference" gaged scenario (modeling baseline conditions using historical land use parameters in a model calibrated to EFCC over a representative climactic period), which is indicated by *.*

	Flow Alteration Metric	Equation	Gaged: *Modeled historical land use	Gaged: Oldest gage record	Ungaged: Nearest neighbor regionalization (EFCC)	Ungaged: Nearest neighbor regionalization (NSE)	Ungaged: SR-HMC
RBI	Flow Alteration (Baseline)	$\frac{(\text{Altered} - \text{Baseline})}{\text{Baseline}}$	-0.2	**0.4**	**0.1**	**0.1**	-0.2
	Flow Alteration (Altered)	$\frac{(\text{Altered} - \text{Baseline})}{\text{Altered}}$	-0.2	**0.4**	**0.1**	**0.1**	-0.2
	Flow Alteration (Average)	$\frac{(\text{Altered} - \text{Baseline})}{\text{Avg}(\text{Altered}, \text{Baseline})}$	-0.2	**0.4**	**0.1**	**0.1**	-0.2
	Alteration Ratio	$\frac{\text{Altered}}{\text{Baseline}}$	-0.2	**0.4**	**0.1**	**0.1**	-0.2
	Alteration	$\text{Altered} - \text{Baseline}$	-0.1	**0.3**	-0.2	**0.1**	-0.2
< 1 cfs	Flow Alteration (Baseline)	$\frac{(\text{Altered} - \text{Baseline})}{\text{Baseline}}$	0.3	0	0.2	0.2	0.2
	Flow Alteration (Altered)	$\frac{(\text{Altered} - \text{Baseline})}{\text{Altered}}$	0.3	0	0.2	0.2	0.2
	Flow Alteration (Average)	$\frac{(\text{Altered} - \text{Baseline})}{\text{Avg}(\text{Altered}, \text{Baseline})}$	0.3	0	0.2	0.2	0.2
	Alteration Ratio	$\frac{\text{Altered}}{\text{Baseline}}$	0.3	0	0.2	0.2	0.2
	Alteration	$\text{Altered} - \text{Baseline}$	0.3	0	0.2	0.2	0.2

4.4 Discussion

4.4.1 Flow-ecology relationships in ungaged basins

Hydrologic modeling of ungaged basins is a crucial component of the ELOHA framework for developing flow-ecology relationships (Poff *et al*., 2010), yet it involves significant uncertainty with no preferred method (Blöschl *et al*., 2013). Neither streamflow flashiness nor drying are particularly easy environmental streamflow metrics to model (Carlisle *et al*., 2010a), especially in semi-arid urbanizing ungaged basins typical of So. CA (Kennard *et al.*, 2010; Nathan and McMahon, 1992; Razavi and Coulibaly, 2017). Despite these obstacles, this study produced accurate flow-ecology relationships in five ungaged basins across the heterogeneous region of So. CA.

The general effects of urbanization on streamflow flashiness and associated flow-ecology relationships are fairly straightforward, with an increase in flashiness causing a decline in ecological condition after urbanization (Paul and Meyer, 2001; Walsh *et al*., 2005; Walsh *et al*., 2001). This inverse flow-ecology relationship was not consistent across ungaged regionalization approaches. Only SR-HMC replicated the inverse relationship (most Spearman correlation coefficients = -0.2, Table 4.6). SR-HMC likely produced appropriate flashiness flow-ecology relationships when both nearest neighbor approaches did not due to its emphasis on regional environmental flow accuracy using statistical techniques including jackknife resampling, clustering, and ensemble averaging. While flow-ecology relationships were only developed at five sites, e complexity of watersheds in So. CA may inhibit spatial similarity of model parameters, alone, for estimating streamflow flashiness and associated flow-ecology relationships in ungaged basins. Developing SR-HMC for high accuracy in a heterogenous region (Ch. 3) may have played a role in its successful modeling of flashiness flow-ecology relationships in ungaged basins.

Flow-ecology relationships between CSCI and different stream drying alteration metrics were similarly modeled in ungaged basins using all three of the regionalization approaches (Spearman correlation coefficient = 0.3 gaged vs. 0.2 ungaged, Table 4.6). Choice of regionalization did not majorly affect stream drying flow-ecology relationships. However, the decline in ecological condition associated with the collective increase in streamflow flashiness and permanence after urbanization found by Roy *et al*., (2005) was only modeled in ungaged basins using SR-HMC.

The three ungaged estimates of flow alteration utilized representative precipitation data from WY 2005-2007. This and previous studies in So. CA have prioritized representative dry, average, and wet years; however, these general regional climate trends do not capture the temporal and sub-regional precipitation patterns that strongly influence streamflow flashiness (Baker *et al.*, 2004). As such, altered and baseline RBI, and their associated flashiness flow-ecology relationships, are more difficult to model than those for streamflow drying. Unlike with streamflow flashiness, the general regional climate trends captured during this period appear to be stronger drivers of streamflow drying than the differences in model parameters generated by different regionalization techniques. This overall control of precipitation on flow permanence likely led to the similar flow-ecology relationships between gaged and ungaged sites for representative climate conditions. The findings of this study wherein stabilization and augmentation of extremely low flows, created by urbanization, lend to wetter conditions that negatively affect ecological condition are consistent with the most frequently performed flow-ecology studies (Poff and Zimmerman, 2010); however, the relationship between urbanization and streamflow drying can vary, with some studies showing dryer conditions after urbanization (Konrad and Booth, 2005; Poff and Zimmerman, 2010; Walsh *et al*., 2005). This discrepancy in the effects of urbanization on flow permanence and ecology is controlled not only by natural factors, such as regional climate and topography, but also engineered water supply and treatment infrastructure and management practices. As such, local information regarding the effects of urbanization on flow permanence should be used to assess flow-ecology relationships. Two independent studies in So. CA have shown increases in permanent water after urbanization, and associated decreases in ecological condition (Riley *et al*., 2005; White and Greer, 2006), lending credibility to the flow-ecology relationships with < 1 cfs in this study.

Ultimately, modeling environmental streamflow, and associated flow-ecology relationships, in ungaged basins is difficult. Results from this study show how choice in regionalization approach can significantly affect the accuracy of ungaged flow-ecology relationships. Despite the strong performance of SR-HMC for flow-ecology relationships with RBI, it is a resource intensive approach to PUB, and not always necessary for accurate modeling. Statistical methods for regionalizing ungaged basins (Carlisle *et* *al.*, 2010a; Sanborn and Bledsoe, 2006) include USGS regional regression equations, and are useful in hydrologically homogeneous regions, especially when data are scarce. In general, SR-HMC could be used to estimate ungaged streamflows and flow-ecology relationships in highly complex regions with heterogeneous hydrogeology and/or flow regimes. Moderate to high data availability is needed for this approach. In contrast, nearest neighbor regionalization is best applied in dense, relatively homogeneous data networks.

4.4.2 Quantifying flow alteration for flow-ecology relationships

Overall, flow-ecology relationships were not very sensitive to formulation of flow alteration, regardless of gaged or ungaged approach. For each modeling approach, streamflow drying produced identical flow-ecology relationships across all five flow alteration metrics (Table 4.6). For the five sites tested in this study, flow-ecology relationships with < 1 cfs were not sensitive to flow alteration metric or regionalization method, provided a period of regionally representative precipitation was modeled. In contrast, flashiness flow-ecology relationships were slightly more sensitive to flow alteration metric than stream drying, with four out of five metrics of flow alteration producing identical flow-ecology relationships (Table 4.6). These less consistent relationships can likely be attributed to the difficulty modeling RBI, whether historical or altered. Results from this study suggest formulation of flow alteration does not have a major impact on flow-ecology relationships. In general, flow-ecology relationships with RBI were slightly more sensitive to flow alteration metric than relationships with < 1 cfs, and were much more sensitive to ungaged regionalization approach.

While the four metrics normalizing flow alteration produced consistent flow-ecology relationships, Alteration Ratio (as $\frac{\text{Current}}{\text{Historic}}$, Table 4.1) contains some ideal properties not shared by the three variations of Flow Alteration. Whereas values of Alteration Ratio center around one, Flow Alteration can be extremely small or large if individual values or differences between baseline and altered conditions are small, even if it is applied as a percent. These sensitive quantities of flow alteration are more difficult to interpret than those centered around one. Another major benefit of Alteration Ratio for generating flow-ecology relationships is its similarity with the $\frac{\text{Observed}}{\text{Expected}}$, or $\frac{\text{O}}{\text{E}}$, bioassessment metric (Hawkins, 2006). In essence, Alteration Ratio and $\frac{\text{O}}{\text{E}}$ are the same metric applied to two different measures of alteration: environmental streamflow condition and taxonomic completeness. Paring $\frac{\text{O}}{\text{E}}$ streamflow and bioassessment metrics has helped provide intuitive flow-ecology relationships across the contiguous United States (Carlisle *et al*., 2010a). Flow-ecology relationships in CA using CSCI, of which $\frac{\text{O}}{\text{E}}$ is a component, can benefit from similarly intuitive relationships when Alteration Ratio is used.

While Alteration Ratio provides clear benefits as a measure of flow alteration, it is not ideal for every flow metric. Deviations between pre- and post-developed flow metrics may be better characterized by % Alteration and metrics characterizing timing, duration, and frequency may be easily analyzed using simple alteration. As such,% Flow Alteration has been used extensively in the environmental streamflow literature (Buchanan *et al.*, 2013; Kennen *et al.*, 2013; McManamay *et al.*, 2013; Poff and Zimmerman, 2010). The traditional application of % Flow Alteration has involved normalizing the difference between altered and baseline streamflow by baseline conditions (Table 4.1). From a modeling perspective, especially in ungaged basins, normalizing by baseline conditions introduces more uncertainty than normalizing by current conditions because fewer streamflow gages are available with accurate long-term records of pre-altered streamflow conditions than gages with more recent flows. With fewer gages, less data are available to build, calibrate, and regionalize models. Results from this study of five sites show the same flow-ecology relationships using Flow Alteration normalized by baseline, altered, and averaging baseline and altered conditions, regardless of gaged or ungaged approach for generating flow alteration (Table 4.6). Based on this finding, normalizing % Flow

Alteration by altered conditions is an option that should be explored further for improving the most commonly used metric of flow alteration from a modeling uncertainty perspective.

4.4.3 Estimating pre-altered baseline streamflow for flow alteration

Estimating baseline streamflow is necessary for quantifying flow alteration and creating flow-ecology relationships. In this study, estimating baseline streamflow at gage sites by adjusting model parameters to reflect historical conditions was considered more accurate than using the oldest available gage data. This method of changing rainfall-runoff model parameters associated with developed land use is consistent with other ELOHA studies. In the Potomac River Basin, historical land use and catchment data was used in conjunction with Hydrological Simulation Program—Fortran (HSPF) to alter calibrated parameters to represent historical conditions (Buchanan *et al.*, 2013). A similar approach has been explored in CA relating curve number to a net initial loss parameter (Sengupta *et al.*, 2018). In New Jersey, a combined approach was used in which baseline conditions were estimated using historical gage data and a state-wide hydrological model, while current flow conditions were estimated from current gage data (Kennen *et al.*, 2013). Water accounting parameters for reservoir storage, diversions, and flow returns have been used to estimate historical conditions using the Colorado's State Water Supply Model (State of Colorado, 2016).

In this study, the historical modeling approach produced slightly less permanent and less flashy baseline conditions compared to altered conditions (Table 4.4). These historical flow conditions were consistently modeled by all five sites for both RBI and < 1 cfs. Less flashy baseline hydrographs can be attributed to the increase in Time of Concentration associated with a longer flow path and no imperviousness, while the increase in flow permanence is likely due to the increase of "urban slobber" (Wolch, 2007). The magnitude of the difference between pre- and post-altered streamflow flashiness and drying were heavily correlated with imperviousness (RBI Spearman ρ = -0.89; < 1 cfs Spearman ρ = 0.69), indicating how watersheds undergoing the most urbanization show the largest increases in streamflow flashiness and permanence for altered conditions. These findings of flashier and more permanent hydrographs with increasing urbanization are consistent with the relevant literature (Paul and Meyer, 2001; Poff and Zimmerman, 2010; Riley *et al*., 2005; Walsh *et al*. 2005; Walsh *et al*., 2001; White and Greer, 2006) and instill confidence in the methods of regionalizing historical streamflow.

4.4.4 Environmental streamflow management in ungaged basins

Bioassessment sites used for flow-ecology relationships are often located along streams with no reliable and/or representative streamflow gage data (Poff *et al*., 2010). Environmental flow management in these ungaged basins is a significant challenge because models must be used to estimate both altered and baseline streamflow conditions, without the aid of calibration data. Based on this study, and reinforced by previous work in ungaged So. CA basins (Ch. 2; Ch. 3), a variety of generally applicable recommendations are made for modeling and environmental flow management in ungaged watersheds.

First, regardless of method for PUB, calibration criteria should focus on the elements of the environmental flow regime most amenable to management. Identifying elements of the flow regime for management is a useful initial step to help tighten model accuracies and streamline studies. Selecting environmental flow components can involve a literature review and/or an initial study of ecological sensitivity. Both were considered in So. CA before selecting streamflow flashiness and permanence (Gasith and Resh, 1999; Mazor *et al*., 2018; Parker *et al*., 2019)

Second, because a variety of methods with varying complexities exist for regionalizing streamflow in ungaged basins, and all environmental streamflow studies are different, strong consideration should be made for a variety of modeling approaches.

Third, multimetric bioassessment indices reduce some challenges with ungaged watershed modeling. Because multimetric indices compare observed ecological condition to regional reference expectations, they are not reliant on two sampling dates to quantify decline (or improvement) of ecological condition. As such, ungaged watershed models can be created for representative precipitation conditions, preferably during a current time period representative of altered streamflow, instead of for two different time periods when ecological sampling might have occurred if a multimetric index was not used. This greatly reduces modeling uncertainty and increases confidence in the results of uncalibrated models.

Fourth, Alteration Ratio (Table 4.1) is a good measure of flow alteration for many flow metrics. It is easy to interpret and robust against flow metrics with small values. Furthermore, Alteration Ratio provides continuity in flow-ecology relationships when paired with $\frac{\text{Observed}}{\text{Expected}}$ ecological metrics, or a multimetric index containing $\frac{\text{O}}{\text{E}}$.

Fifth and finally, when choosing elements of the flow regime for developing environmental flow criteria, modeling judgement on the ability to accurately estimate specific baseline and altered metrics in ungaged basins should be heavily considered. Opportunities to produce poor estimates of ungaged streamflow conditions should be avoided to increase confidence in the results of regionalized models. Focusing on elements of the environmental flow regime most amenable to management requires not only considering their ecological significance and manageability, but also their propensity to be modeled accurately.

For the five case study sites modeled as ungaged basins in So. CA, streamflow permanence (as < 1 cfs) was significantly easier to model for altered, and historical conditions (Table 4.4; Table 4.5) than flashiness (as RBI). Subsequently, flow-ecology relationships were stronger and more robust when the multimetric bioassessment index, CSCI, was paired with measures of < 1 cfs alteration (Table 4.6). Flashiness flow-ecology relationships were inconsistent (Table 4.6). A potential benefit of focusing environmental flow management in So. CA on flow permanence is its relative annual resilience to precipitation. Because precipitation is the primary input of rainfall-runoff models commonly used to regionalize ungaged basins, modeling confidence of < 1 cfs can be relatively high across different historical and current time periods. Finally, while Alteration Ratio is recommended as an initial measure of flow alteration, flow-ecology relationships with permanence were less sensitive to different metrics than those with flashiness (Table 4.6). As such, < 1 cfs flow-ecology relationships in So. CA are not sensitive to how flow alteration is quantified.

4.5 Conclusions

Accurately quantifying flow alteration at ungaged sites is necessary to produce reliable flow-ecology relationships for effective environmental streamflow management. Through this study, methods of regionalizing ungaged watersheds and formulating flow alteration were identified to produce effective flow-ecology relationships. While the specific results regarding flow permanence and flashiness are regional to So. CA and the series of five case study sites, this study produced general guidelines for improving environmental streamflow management in ungaged watersheds. Streamflow Regionalization with Hydrologic Model-based Classification (Ch. 3) has utility for modeling ungaged baseline and altered streamflow conditions in highly heterogeneous regions, but it is relatively resource intensive and may not always be required for certain model applications. However, linking this regionalization method with a series of recommended PUB practices can generally provide accurate and intuitive flow-ecology relationships. These modeling practices include using management-amenable environmental flow metrics, ecologically-relevant calibration of rainfall-runoff models, a multimetric bioassessment index, and Alteration Ratio for quantifying most metrics of flow alteration.

This study assessed the impacts of modeling and flow alteration choices on flow-ecology relationships in ungaged basins. As such, a full analysis of So. CA flow-ecology relationships was outside the scope. Future work with environmental flow management in So. CA should focus on developing complex flow-ecology relationships, starting with flow permanence, and using more sites. As done in previous studies around the country, quantile (Buchanan *et al.*, 2013; Konrad *et al.*, 2008) or linear regression (Kennen *et al.*, 2010; McManamay *et al.*, 2013; Pomeroy *et al.*, 2008) can quantify ecological condition for different values of flow alteration. As opposed to the general correlation relationships of this study, regression-based flow-ecology relationships can be used to directly produce practical environmental flow targets. As a part of this expanded analysis, new bioassessment sites could be developed near streamflow gages with data older than WY 1989. Ideally, selecting sites with data from at least 1965 would allow for further assessing baseline flow conditions in accordance with the timing of urbanization (White and Greer, 2006).

The methods supporting environmental flow development with ELOHA have steadily improved over the past ten years, but boundaries can always be pushed. The superiority of multimetric bioassessment indices suggests potential for a type of multimetric index of flow alteration. Like CSCI, a multimetric index of flow alteration would compare modeled or gaged data to regional standards of the flow regime. As it is refined, such a multimetric index would contain weighted components of different flow regime elements. These weights could be managed to provide one simple value of the most dynamically relevant environmental flow alteration to pair with one measure of ecological change. Such an approach might simplify environmental flow management and facilitate its continued implementation. The recommended ungaged environmental flow management techniques identified in So. CA, including SR-HMC, choosing environmental flow metrics amenable to management, calibrating rainfall-runoff models to ecologically-relevant criteria, using a multimetric bioassessment index, and quantifying flow alteration with Alteration Ratio when feasible, should be extended to other regions. Considering the ubiquity of bioassessment sites quantifying CSCI in California, another region of CA would be a natural place to extend the methods of this and previous studies (Ch. 2; Ch. 3). As improvements are made to the methods supporting environmental flow management, policies protecting streams from hydromodification will grow and help conserve the ecological integrity and designated uses of streams.

Chapter 5

Conclusions

Streamflow prediction in ungaged basins is a difficult task needed for a wide range of hydrologic and hydraulic analyses. These difficulties are amplified when developing environmental streamflow criteria due to the challenges with estimating environmental flow metrics. Modeling ungaged basins is further complicated in heterogeneous regions with diverse land use, geologic settings, hydroclimatological processes, and streamflow regimes, where neighboring watersheds may differ substantially. Scientists and engineers must look to address these issues by improving techniques for PUB. Doing so will strengthen the scientific integrity and confidence of analyses. This is especially important in the growing practice of environmental streamflow management, where critical bioassessment sites are typically located in ungaged basins. My dissertation is built around addressing some important questions affecting the management of water resources and environment streamflows in ungaged basins. Specifically, can uncertain parameters in models of ungaged basins be estimated with more confidence? And, does reducing uncertainty in these parameters produce more accurate and consistent estimates of streamflow, flow alteration, and flow-ecology relationships in ungaged basins from a heterogeneous region?

These are the two primary questions of my dissertation, which I first address in Chapter 2 with a new method for classifying streams called “Hydrologic Model-based Classification” (HMC). Grouping hydrologically similar streams according to the regional accuracy of calibrated rainfall-runoff parameter is a new way of thinking about hydrologic similarity. Parallels between the regional nature of ELOHA and regionalization of ungaged basins facilitated the idea reevaluate hydrologic similarity from the perspectives of hydrologic modeling, ungaged basins, and reducing parameter uncertainty. . In water resources management, modeling ungaged basins is often overshadowed by gage data but it becomes crucial when such data are unreliable or unavailable. The accuracy of hydrologic models should be considered as a component of stream classification when the analysis might facilitate regionalization of ungaged basins. In HMC, jackknife resampling is used to transfer intact calibrated parameter sets from each model to every other model from a regional catalog. This process creates an error matrix describing the regional transferability of calibrated parameters. After grouping models with reciprocally accurate parameter sets, multinomial logistic regression with watershed characteristics can assign a new ungaged site to a group of parameter sets likely to produce accurate results in a model of that ungaged basin. Reducing the pool of infinitely possible parameter sets, down to a small regionally-tested group likely to perform well in the model of an ungaged basin, greatly reduces parameter uncertainty.

In Chapter 3, I expand HMC with ensemble rainfall-runoff modeling of geographically weighted model output to produce a new regionalization framework for ungaged basins, called “Streamflow Regionalization with Hydrologic Model-based Classification” (SR-HMC). While typical nearest-neighbor regionalization reduces some parameter uncertainty in ungaged basins, this uncertainty is constrained by the hydrologic similarly between the ungaged location and nearest calibrated model. In general, geographic proximity is a good measure of hydrologic similarity, but it loses strength in diverse regions or sparse networks. My modeling framework combines the general accuracy of geographic proximity with the power of HMC and model output averaging to improve streamflow estimates at ungaged locations across a heterogeneous landscape by reducing parameter uncertainty, compared to nearest-neighbor regionalization. This framework was developed for any regional analysis of ungaged basins, not just those related to environmental streamflow, but I demonstrate SR-HMC with environmental streamflow metrics in heterogeneous coastal southern California.

In Chapter 4, I focus on management endpoints of environmental flows in ungaged basins. I recommend modeling practices and decisions that can increase the accuracy and reliability of flow alteration and flow-ecology relationships in ungaged basins. I further recommend how to identify ideal streamflow metrics robust against ungaged regionalization modeling method, flow alteration metric, and time period-specific precipitation inputs. My modeling innovations and recommendations have been demonstrated to increase the accuracy and reduce the uncertainty of streamflow, flow alteration, and flow-ecology relationships in ungaged basins across a heterogeneous region.

While HMC shows tremendous potential for reducing parameter uncertainty in models of ungaged basins, the statistical analyses of my dissertation were performed with a relatively small sample size (5 validation sites and 25 ensemble sites). Studying environmental streamflow in a region as heterogenous and modified as So. CA greatly limited the number of sites. As such, results indicating the increased accuracy of HMC and SR-HMC over traditional classifications and nearest neighbor regionalization are not definitive and should be tested in other regions, ideally with larger sample sizes. A natural extension to regions further north of Santa Barbara would provide valuable information regarding environmental streamflow management and modeling ungaged basins in California. However, looking forward, HMC and SR-HMC should be expanded to brand new regions, not only for environmental flows but also other applications. New systems with limited hydrologic diversity and dense gage networks would provide a deeper understanding of regional parameter uncertainty and test the utility of the methods developed in my dissertation. Focusing modeling efforts on simple metrics in such regions, such as overall fit or peak flow, would provide a clearer assessment of HMC and SR-HMC.

My dissertation builds on previous advances in modeling ungaged basins and environmental streamflow management to introduce and test new ideas. I develop a novel framework (SR-HMC) for modeling ungaged basins that reduces parameter uncertainty. In developing SR-HMC, I test new approaches within the ELOHA framework and advance the analysis of streamflow flashiness and drying as management endpoints in So. CA. Prioritizing the ecological integrity of streams by managing them in harmony with their natural flow regimes provides a path towards sustainable freshwater resources. My dissertation provides technical tools to help facilitate this future.

References

Allan, J.D, 2004. Landscapes and Riverscapes: The Influence of Land Use on Stream Ecosystems. *Annual Review of Ecology, Evolution, and Systematics* 35:257-284.

Angradi, T.R., M.S. Pearson, D.W. Bolgrien, T.M. Jicha, D.L. Taylor, B.H. Hill, 2009. Multimetric macroinvertebrate indices for mid-continent US great rivers. *Freshwater Science* 28(4):785-804.

Arizona Department of Transportation (ADOT), 2014. Highway Drainage Design Manual, Volume 2: Hydrology. Phoenix, AZ, 182 p., URL: https://azdot.gov/sites/default/files/2019/07/2014_adot_hydrology_manual.pdf.

Armanini, D.G., N. Horrigan, W.A. Monk, D.L. Peters, and D.J. Baird, 2011. Development of a benthic macroinvertebrate flow sensitivity index for Canadian rivers. *River Research and Applications* 27(6):723-737.

Armstrong, D.S., T.A. Richards, and S.B. Levin, 2011. Factors Influencing Riverine Fish Assemblages in Massachusetts. Scientific Investigations Report 2011-5193, U.S. Department of the Interior, U.S. Geological Survey (USGS), 74 p., URL: https://pubs.usgs.gov/sir/2011/5193/pdf/sir2011-5193_text_508.pdf.

Arnold, C.L. and J. Gibbons, 1996. Imperious Surface Coverage: The Emergence of a Key Environmental Indicator. *Journal of American Planning Association* 62(2):243-258.

Arsenault, R., M. Breton-Dufour, A. Poulin, G. Dallaire, and R. Romero-Lopez, 2019. Streamflow prediction in ungauged basins: analysis of regionalization methods in a hydrologically heterogeneous region of Mexico. *Hydrological Sciences Journal* 64(11):1297-1311.

Arsenault, R. and F.P. Brissette, 2014. Continuous streamflow prediction in ungauged basins: The effects of equifinality and parameter set selection on uncertainty in regionaliztion approaches. *Water Resources Research* 50(7):6135-6153.

Arthington, A.H., A. Bhaduri, S.E. Bunn, S.E. Jackson, R.E. Tharme, D. Tickner, B. Young, M. Acreman, N. Baker, S. Capon, A.C. Horne, E. Kendy, M.E. McClain, N.L. Poff, B.D. Richter, and S. Ward, 2018. The Brisbane Declaration and Global Action Agenda on Environmental Flows. *Frontiers in Environmental Science* 6(45):1-15.

Arthington, A.H., S.E. Bunn, N.L. Poff, and R.J. Naiman, 2006. The Challenge of Providing Environmental Flow Rules to Sustain River Ecosystems. *Ecological Applications* 16(4):1311-1318.

Baker, D.B., R.P. Richards, T.T. Loftus, and J.W. Kramer, 2004. A New Flashiness Index: Characteristics and Applications to Midwestern Reivers and Stream. *Journal of the American Water Resources Association* 40(2):503-522.

Barbour, M.T., J. Gerritsen, B.D. Snyder, and J.B. Stribling, 1999. Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish, Second Edition. EPA 841-B-99-002. United States Environmental Protection Agency (EPA) Office of

Water, Washinton, D.C. URL: https://www3.epa.gov/region1/npdes/merrimackstation/pdfs/ar/AR-1164.pdf.

Barbour, M.T., J.B. Stribling, and J.R. Karr, 1995. Multimetric approach for establishing biocriteria and measuring biological condition. *In* W.S. Davis and T.P. Simon (editors) Biological assessment and criteria: tools for water resource planning and decision making. Lewis Publishers, Boca Raton, Florida.

Belmar, O., J. Velasco, and F. Martinez-Capel, 2011. Hydrologic Classification of Natural Flow Regimes to Support Environmental Flow Assessments in Intensively Regulated Mediterranean Rivers, Segura River Basin (Spain). *Environmental Management* 47(5):992-1004.

Bardossy, A., 2007. Calibration of hydrological model parameters for ungauged matchments. *Hydrology and Earth System Sciences* 11(2):703-710.

Beven, K.J., 1989. Changing ideas in hydrology – the case of physically-based models. *Journal of Hydrology* 105(1-2):157-172.

Beven, K.J., 2012. Rainfall-Runoff Modelling: The Primer – Second Edition. John Wiley & Sons, West Sussex, United Kingdom.

Beven, K. and A. Binley, 1992. The Future of Distributed Models: Model Calibration and Uncertainty Prediction. *Hydrological Processes* 6:279-298.

Bledsoe, B.P., M.C. Brown, and D.A. Raff, 2007. GeoTools: A Toolkit for Fluvial System Analysis. *Journal of the American Water Resources Association* 43(3):757-772.

Blöschl, G., M. Sivapalan, T. Wagener, A. Viglone, and H. Savenije, 2013. Runoff Prediction in Ungauged Basins: Synthesis across Processes, Places, and Scales. Cambridge University Press, London, United Kingdom.

Blume, T., E. Zehe, and A. Bronstert, 2007. Rainfall-runoff response, event-based runoff coefficients and hydrograph separation. *Hydrologic Sciences Journal* 52(5):843-862.

Booker, D.J. and R.A. Woods, 2014. Comparing and combining physically-based and empirically-based approaches for estimating hydrology of ungauged catchments. *Journal of Hydrology* 508:227-239.

Booth, D.B., J.R. Karr, S. Schauman, C.P. Konrad, S.A. Morley, M.G. Larson, and S.J. Burges, 2004. Reviving Urban Streams: Land Use, Hydrology, Biology, and Human Behavior. *Journal of the American Water Resources Association*, 40(5):1351–1364.

Brown, S.C., R.E. Lester, V.L. Versace, J. Fawcett, and L. Laurenson, 2014. Hydrologic Landscape Regionalisation Using Deductive Classification and Random Forest. *PLoS One* 9(11):e112856.

Buchanan, C., K. Foreman, J. Johnson, and A. Griggs, 2011. Development of a Basin-wide Benthic Index of Biotic Integrity for Non-tidal Streams and Wadeable Rivers in the Chesapeake Bay Watershed: Final Report to the Chesapeake Bay Program Non-Tidal Water Quality Workgroup. Interstate

Commission on the Potomac River Basin, Rockville, Maryland, 75 p. URL: http://www.potomacriver.org/wp-content/uploads/2014/12/ICPRB11-011.pdf.

Buchanan, C., H.L.N. Moltz, H.C. Haywood, J.B. Palmer, and A.N. Griggs, 2013. A test of The Ecological Limits of Hydrologic Alteration (ELOHA) method for determining environmental flows in the Potomac River basin, U.S.A. *Freshwater Biology* 58:2632-2647.

Bunn, S.E. and A.H. Arthington, 2002. Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. *Environmental Management* 30(4):492–507.

Burn, D.H. and D.B. Boorman, 1993. Estimation of hydrologic parameters at ungauged catchments. *Journal of Hydrology* 143(3-4):429-454.

Carlisle, D.M., J. Falcone, D.M. Wolock, M.R. Meador, and R.H. Norris, 2010a. Predicting the natural flow regime: models for assessing hydrological alteration in streams. *River Research Applications* 26: 118–36.

Carlisle, D.M., S.M. Nelson, and J. May, 2016. Associations of stream health with altered flow and water temperature in the Sierra Nevada, California. *Ecohydrology* 9:930-941.

Carlisle, D.M., D.M. Wolock, and M.R. Meador, 2010b. Alteration of streamflow magnitudes and potential ecological consequences: a multiregional assessment. *Frontiers in Ecology and the Environment* 9(5):264-270.

Carpenter, T.M. and K.P. Georgakakos, 2006. Intercomparison of lumped versus distributed hydrologic model ensemble simulations on operational forecast scales. *Journal of Hydrology* 329(1-2):174-185.

Cassin, J., R. Fuerstenberg, L. Tear, K. Whiting, D. St. John, B. Murray, and J. Burkey, 2005. Development of Hydrological and Biological Indicators of Flow Alteration in Puget Sound Lowland Streams. King County Water and Land Resources Division Final Report, Seattle, Washington, 90 p.

Castellarin, A., G. Galeati, L. Brandimarte, A. Montanari, and A. Brath, 2004. Regional flow-duration curves: reliability for ungauged basins. *Advances in Water Resources* 27(10):953-965.

Castiglioni, S., A. Castellarin, and A. Montanari, 2009. Prediction of low-flow indices in ungauged basins through physiographical space-based interpolation. *Journal of Hydrology* 378:272-280.

Chang, F., W. Qiu, R.H. Zamar, R. Lazarus, and X. Wang, 2010. clues: An R Package for Nonparametric Clustering Based on Local Shrinking. *Journal of Statistical Software* 33(4):1-16.

Charrad, M., N. Ghazzali, V. Boiteau, and A. Niknafs, 2014. NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set. *Journal of Statistical Software* 61(6):1-36.

Chen, W., and J.D. Olden, 2017. Evaluating transferability of flow-ecology relationships across space, time, and taxonomy. *Freshwater Biology* 63(8):817-830.

Chery Jr., D.L., 1966. Design and Test of a Physical Watershed Model. *Journal of Hydrology* 4:224-235.

Chin, A., 2006. Urban transformation of river landscapes in a global context. *Geomorphology* 79(3-4):460-487.

Clausen, B., and B.J.F. Biggs, 1997. Relationships between benthic biota and hydrological indices in New Zealand streams. *Freshwater Biology* 38(2):327-342.

Crawford, N.H. and R.K. Linsley, 1966. Digital Simulations in Hydrology: Stanford Watershed Model IV. Stanford University Department of Civil Engineering Technical Report no. 39, 210 p.

Davies, P.M., R.J. Naiman, D.M. Warfe, N.E. Pettit, A.H. Arthington, and S.E. Bunn, 2014. Flow-ecology relationships: closing the loop on effective environmental flows. *Marine and Freshwater Research* Short Communication 65:133-141.

Duan, Q., N.K. Ajami, X. Gao, and S. Sorooshian, 2007. Multi-model ensemble hydrologic prediction using Bayesian model averaging. *Advances in Water Resources* 30:1371-1386.

Duan, Q., J. Schaake, V. Andreassian, S. Franks, G. Goteti, H.V. Gupta, Y.M. Gusev, F. Habets, A. Hall, L. Hay, T. Hogue, M. Huang, G. Leavesley, X. Liang, O.N. Nasonova, J. Noilhan, L. Oudin, S. Sorooshian, T. Wagener, and E.F. Wood, 2006. Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshop. *Journal of Hydrology* 320(1-2):3-17.

Duan, Q., S. Sorooshian, and V. Gupta, 1992. Effective and Efficient Global Optimization for Conceptual Rainfall-Runoff Models. *Water Resources Research* 28(4):1015-1031.

Eddy, M.C., J. Phelan, L. Patterson, J. Allen, and S. Pearsall, 2017. Evaluating low metric-based stream classification systems to support the determination of ecological flows in North Carolina. *Journal of the American Water Resources Association* 53(1):30-41.

Efron, B., 1982. The Jackknife, the Bootstrap, and other Resampling Plans. CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

Falcone, J., 2011. GAGES-II: Geospatial Attributes of Gauges for Evaluating Streamflow. U. S. Department of the Interior, U.S. Geological Survey, URL: https://water.usgs.gov/GIS/metadata/usgswrd/XML/gaugesII_Sept2011.xml.

Farmer, W.H. and R.M. Vogel, 2013. Performance-weighted methods for estimating monthly streamflow at ungauged sites. *Journal of Hydrology* 477:240-250.

Falkenmark, M., J. Lundqvist, and C. Widstrand, 1989. Macro-scale water scarcity requires micro-scale approaches: Aspects of vulnerability in semi-arid development. *Natural Resources Forum* 13(4):258-267.

Franchini, M., and M. Pacciani, 1991. Comparative analysis of several conceptual rainfall-runoff models. *Journal of Hydrology* 122(1-4):161-219.

Freeman, M.C., and P.A. Marcinek, 2006. Fish Assemblage Responses to Water Withdrawals and Water Supply Reservoirs in Piedmont Streams. *Environmental Management* 38(3):435-450.

Friedl, H. and E. Stampfer, 2014. Jackknife Resampling. Wiley StatsRef: Statistics Reference Online.

Fry, J., G. Xian, S. Jin, J. Dewitz, C. Homer, L. Yang, C. Barnes, N. Herold, and J. Wickham, 2011. Completion of the 2006 National Land Cover Database for the conterminous United States. *Photogrammetric Engineering and Remote Sensing* 77(9):858-863.

Gasith, A. and V.H. Resh, 1999. Streams in Mediterranean Climate Regions: Abiotic Influences and Biotic Responses to Predictable Seasonal Events. *Annual Review of Ecology, Evolution, and Systematics* 30:51-81.

Gelman, A., J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari, and D.B. Rubin, 2013. Bayesian Data Analysis, Third Edition. Chapman and Hall, New York, NY.

Gotvald, A.J., N.A. Barth, A.G. Veilleux, and C. Parrett, 2012. Methods for Determining Magnitude and Frequency of Floods in California, Based on Data through Water Year 2006. U.S. Geological Survey Scientific Investigations Report 2012-5113, 38p., URL: https://pubs.usgs.gov/sir/2012/5113/pdf/sir2012-5113.pdf.

Grayson, R.B., I.D. Moore, and T.A. McMahon, 1992. Physically Based Hydrologic Modeling 2. Is the Concept Realistic? *Water Resources Research* 26(10): 2659-2666.

Hawley, R.J. and B.P. Bledsoe, 2011. How do flow peaks and durations change in suburbanizing semi-arid watersheds? A southern California case study. *Journal of Hydrology* 405(1-2):69-82.

Hawley, R.J. and B.P. Bledsoe, 2013. Channel enlargement in semiarid urbanizing watersheds: A southern California case study. *Journal of Hydrology* 496:17-30.

Hawkins, C.P., 2006. Quantifying Biological Integrity by Taxonomic Completeness: Its Utility in Regional and Global Assessments. *Ecological Applications* 16(4):1277-1294.

Hawkins, C.P. and M.R. Vinson, 2000. Weak correspondence between landscape classifications and stream invertebrate assemblages: implications for bioassessment. *Journal of the North American Benthological Society* 19(3):501-517.

Helsel, D.R. and R.M. Hirsch, 2002. Statistical Methods in Water Resources. Techniques of Water Resources Investigations, Book 4, Chapter A3. U.S. Geological Survey (USGS). 522p., URL: https://pubs.usgs.gov/twri/twri4a3/pdf/twri4a3-new.pdf.

Henriksen, J.A., J. Heasley, J.G. Kennen, and S. Nieswand, 2006. Users' Manual for the Hydroecological Integrity Assessment Process Software (Including the New Jersey Assessment Tools). Open-File Report 2006-1093, USGS, Fort Collins Science Center, Fort Collins, CO, 80 p.

Hrachowitz, M., H.H.G. Savenije, G. Blöschl, J.J. McDonnell, M. Sivapalan *et al.*, 2013. A decade of Predictions in Ungauged Basins (PUBS)—a review. *Hydrological Sciences Journal* 58(6):1198-1255.

Hubert, L. and P. Arabie, 1985. Comparing partitions. *Journal of Classification* 2:193-218.

Hydrologic Engineering Center, 2015. Hydrologic Modeling System HEC-HMS User's Manual Version 4.1. U.S. Army Corps of Engineers Report, 600p., URL: http://www.hec.usace.army.mil/software/hec-hms/documentation/HEC-HMS_Users_Manual_4.1.pdf.

Iowa Statewide Urban Design and Specifications (SUDAS), 2013. Design Manual Chapter 2 – Stormwater 2B – Urban Hydrology and Runoff. Iowa State University, Ames, IA, 14p.

Jain, S.K., and K.P. Sudheer, 2008. Fitting of Hydrologic Models: A Close Look at the Nash-Sutcliffe Index. *Journal of Hydrologic Engineering* 13(10):981-986.

Kendy, E., C. Apse, K. Blann, M.P. Smith, and A. Richardson, 2012. A Practical Guide to Environmental Flows for Policy and Planning with Nine Case Studies in the United States. The Nature Conservancy Report, 72p., URL: https://www.conservationgateway.org/ConservationByGeography/NorthAmerica/UnitedStates/edc/Documents/ED_freshwater_envflows_Practical%20Guide%20Eflows%20for%20Policy.pdf.

Kennard, M.J., S.J. MacKay, B.J. Pusey, J.D. Olden, and N. Marsh, 2010. Quantifying Uncertainty in Estimation of Hydrologic Metrics for Ecohydrological Studies. *River Research and Applications* 26(2):137-156.

Kennen, J.G., L.J. Kauffman, M.A. Ayers, D.M. Wolock, and S.J. Colarullo, 2008. Use of an integrated flow model to estimate ecologically relevant hydrologic characteristics at stream biomonitoring sites. *Ecological Modelling* 211(1-2):57-76.

Kennen, J.G., M.L. Riskin, P.A. Reilly, and S.J. Colarullo, 2013. Method to Support Total Maximum Daily Load Development Using Hydrologic Alteration as a Surrogate to Address Aquatic-Life Impairment in New Jersey Streams. Scientific Investigations Report 2013-5089, U.S. Department of the Interior, U.S. Geological Survey, 98 p., URL: https://pubs.usgs.gov/sir/2013/5089/pdf/sir2013-5089.pdf.

Kennen, J.G., K. Riva-Murray, and K.M. Beaulieu, 2010. Determining Hydrologic Factors that Influence Stream Macroinvertebrate Assemblages in the Northeastern US. *Ecohydrology* 3(1):88–106.

Kerans, B.L. and J.R. Karr, 1994. A Benthic Index of Biotic Integrity (B-IBI) for Rivers of the Tennessee Valley. *Ecological Applications* 4(4):768-785.

Kirpich, Z.P., 1940. Time of concentration in small agricultural watersheds. *Civil Engineering* 10(6):362.

Kokkonen, T.S., A.J. Jakeman, P.C. Young, and H.J. Koivusalo, 2003. Prediction daily flows in ungauged catchments: model regionalization from catchment descriptors at the Coweeta Hydrologic Laboratory, North Caroline. *Hydrological Processes* 17(11):2219-2238.

Konrad, C.P. and D.B. Booth, 2005. Hydrologic Changes in Urban Streams and Their Ecological Significance. *American Fisheries Society Symposium* 47:157-177.

Konrad, C.P., A.M. Brasher, and J.T. May, 2008. Assessing Streamflow Characteristics as Limiting Factors on Benthic Invertebrate Assemblages in Streams across the Western United States. *Freshwater Biology*, 53(10):1983–1998.

Kuczera, G. and E. Parent, 1998. Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm. *Journal of Hydrology* 211(1-4):69-85.

Lane, B.A., H.E. Dahlke, G.B. Pasternack, and S. Sandoval-Solis, 2017. Revealing the Diversity of Natural Hydrologic Regimes in California with Relevance for Environmental Flow Applications. 53(2):411-430.

Lane, B.A., S. Sandoval-Solis, E.D. Stein, S.M. Yarnell, G.B. Pasternack, H.E. Dahlke, 2018. Beyond Metrics? The Role of Hydrologic Baseline Archetypes in Environmental Water Management. *Environmental Management* 62:678-693.

Legates, D.R., and G.J. McCabe Jr., 1999. Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water Resources Research* 35(1):233-241.

Li, H., Y. Zhang, F.H.S. Chiew, and S. Xu, 2009. Predicting runoff in ungauged catchments by using Xinanjiang model with MODIS leaf area index. *Journal of Hydrology* 370(1-4):155-162.

Liermann, C.A.R., J.D. Olden, T.J. Beechie, M.J. Kennard, P.B. Skidmore, C.P Konrad, and H. Imaki, 2012. Hydrogeomorphic Classification of Washington State Rivers to Support Emerging Environmental Flow Management Strategies. *River Research and Applications* 28(9):1340-1358.

Liu, Y. and H.V. Gupta, 2007. Uncertainty in hydrologic modeling: Towards an integrated data assimilation framework. *Water Resources Research* 43(7):W07401.

Mackay, S.J., A.H Arthington, and C.S. James, 2014. Classification and comparison of natural and altered flow regimes to support an Australian trial of the Ecological Limits of Hydrologic Alteration framework. *Ecohydrology* 7: 1485–1507.

Marchetti, M.P., and P.B. Moyle, 2001. Effects of Flow Regime on Fish Assemblages in a Regulated California Stream. *Ecological Applications* 11(2):530-539.

Masch, F.D., 1984. Hydrology Design Manual. US Department of Transportation Federal Highway Administration Hydraulic Engineering Circular No. 19 (HEC No. 19, FHWA-1P-84-15), McClean VA, 342 p.

Mazor, R.D., J.T. May, A. Sengupta, K.S. McCune, B.P. Bledsoe, and E.D. Stein, 2018. Tools for managing hydrologic alteration on a regional scale: Setting targets to protect stream health. *Freshwater Biology* 63(8) Special Issue: Evaluating and Managing Environmental Water Regimes in a Water-Scarce and Uncertain Future:786-803.

Mazor, R.D., A.C. Rehn, P.R. Ode, M. Engeln, K.C. Schiff, E.D. Stein, D.J. Gillett, D.B. Herbst, and C.P. Hawkins, 2017. Bioassessment in complex environments: designing an index for consistent meaning in different settings. *Freshwater Science* 35(1):249-271.

McCuen, R.H., 1973. The role of sensitivity analysis in hydrologic modeling. *Journal of Hydrology* 18(1):37-53.

McCuen, R.H., 2005. Accuracy Assessment of Peak Discharge Models. *Journal of Hydrologic Engineering* 10(1):16-22.

McIntyre, N., H. Lee, H. Wheater, A. Young, and T. Wagener, 2005. Ensemble predictions of runoff in ungauged catchments. *Water Resources Research* 41:W12434.

McKay, L., T. Bondelid, T. Dewald, J. Johnston, R. Moore, and A. Rea, 2012. NHDPlus Version 2: User Guide. U.S. Department of the Interior, U.S. Environmental Protection Agency (EPA), URL: https://s3.amazonaws.com/nhdplus/NHDPlusV21/Documentation/NHDPlusV2_User_Guide.pdf.

McManamay, R.A., D.J. Orth, C.A. Dolloff, and E.A. Frimpong, 2012. A regional classification of unregulated stream flows: spatial resolution and hierarchical frameworks. *River Research and Applications* 28(7):1019-1033.

Murphy, J.C., R.R. Knight, W.J. Wolfe, and W.S. Gain, 2013. Predicting Ecological Flow Regime at Ungauged Sites: A Comparison of Methods. *River Research and Applications* 29(5):660-669.

Naiman, R.J., J.J. Magnuson, D.M. McKnight, and J.A. Stanford, 1995. The freshwater imperative: a research agenda. Island Press, Washington DC.

Nash, J.E., and J.V. Sutcliff, 1970. River flow forecasting through conceptual models part I – A discussion of principles. *Journal of Hydrology* 10(3):282-290.

Nathan, R.J., and T.A. McMahon, 1992. Estimating Low Flow Characteristics in Ungauged Catchments. *Water Resources Management* 6(2):85-100.

Noori, N. and L. Kalin, 2016. Coupling SWAT and ANN models for enhanced daily streamflow prediction. *Journal of Hydrology* 533:141-151.

Ode, P.R., A.C. Rehn, and J.T. May, 2005. A Quantitative Tool for Assessing the Integrity of Southern Coastal California Streams. *Environmental Management* 35(4):493-504.

Oksanen, J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P.R. Minchin, R.B. O'Hara, G.L. Simpson, P. Solymos, M. Henry, H. Stevens, E. Szoecs, and H. Wagener, 2019. vegan: Community Ecology Package. R package version 2.5-5. https://CRAN.R-project.org/package=vegan.

Olden, J.D., M.J. Kennard, and B.J. Pusey, 2012. A framework for hydrologic classification with a review of methodologies and applications in ecohydrology. *Ecohydrology* 5(4):503-518.

Ott, R.L. and M. Longnecker, 2008. An Introduction to Statistical Methods and Data Analysis. Brooks/Cole, Cengage Learning, Belmont, CA.

Oudin, L., V. Andreassian, C. Perrin, C. Michel, and N.L. Moine, 2008. Spatial proximity, physical similarity, regression and ungauged catchments: A comparison of regionalization approaches based on 913 French catchments. *Water Resources Research* 44(3):1-15.

Parker, S.R., S.K. Adams, R.W. Lammers, E.D. Stein, and B.P. Bledsoe, 2019. Targeted hydrologic model calibration using ecological flow metrics. *Journal of Hydrology* 573:546-556.

Paul, M.J. and J.L. Meyer, 2001. Streams in the Urban Landscape. *Annual Review of Ecology and Systematics* 32:333-365.

Peel, M.C. and G. Blöschl, 2011. Hydrological modelling in a changing world. *Progress in Physical Geography* 35(2):249-261.

Peñas, F.J., J. Barquín, and C. Álvarez, 2016. Sources of variation in hydrological classifications: Time scale, flow series origin and classification procedure. *Journal of Hydrology* 538:487-499.

Phelan, J., T. Cuffney, L. Patterson, M. Eddy, R. Dykes, S. Pearsall, C. Goudreau, J. Mead, and F. Tarver, 2017. Fish and Invertebrate Flow-Biology Relationships to Support the Determination of Ecological Flows for North Carolina. *Journal of the American Water Resources Association* 53(1):42-55.

Poff, N.L., 1996. A hydrogeography of unregulated streams in the United States and an examination of scale-dependence in some hydrological descriptors. *Freshwater Biology* 36(1):71-91.

Poff, N.L., 1997. Landscape Filters and Species Traits: Towards Mechanistic Understanding and Prediction in Stream Ecology. *Journal of the North American Benthological Society* 16(2):391-409.

Poff, N.L., J.D. Allan, M.B. Bain, J.R. Karr, K.L. Prestegaard, B.D. Richter, R.E. Sparks, and J.C. Stromberg, 1997. The Natural Flow Regime. *BioScience* 47(11):769-784.

Poff, N.L., B.P. Bledsoe, and C.O. Cuhaciyan, 2006. Hydrologic variation with land use across the contiguous United States: Geomorphic and ecological consequences for stream ecosystems. *Geomorphology*, 79(3-4):264-285.

Poff, N.L., B.D. Richter, A.H. Arthington, S.E. Bunn, R.J. Naiman, E. Kendy, M. Acreman, C. Aspe, B.P. Bledsoe, M.C. Freeman, J. Henriksen, R.B. Jacobson, J.G. Kennen, D.M. Merritt, J.H. O'Keeffe, J.D. Olden, K. Rogers, R.E. Tharme, and A. Warner, 2010. The ecological limits of hydrologic alteration (ELOHA): a new framework for developing regional environmental flow standards. *Freshwater Biology* 55(1):147-170.

Poff, N.L. and J.V. Ward, 1989. Implications of Streamflow Variability and Predictability for Lotic Community Structure- A Regional Analysis of Streamflow Patterns. *Canadian Journal of Fisheries and Aquatic Sciences* 46(10):1805-1818.

Pomeroy, C.A., L.A. Roesner, J.C. Coleman, and E. Rankin, 2008. Protocols for Studying Wet Weather Impacts and Urbanization Patterns. The Water Environment Research Foundation (WERF) Final Report, 188 p.

Ponce, V.M. and D.S. Lindquist, 1990. Management of Baseflow Augmentation: A Review. *Water Resources Bulletin* 26(2): 259-268.

Pyne, M.I., D.M. Carlisle, C.P. Konrad, and E.D. Stein, 2017. Classification of California streams using combined deductive and inductive approaches: Setting the foundation for analysis of hydrologic alteration. *Ecohydrology* 10(3):e1802.

Puckridge, J.T., F. Sheldon, K.F. Walker, and A.J. Boulton, 1998. Flow variability and the ecology of large rivers. *Marine Freshwater Research* 49(1):992-1004.

R Core Team, 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

Razavi, T. and P. Coulibaly, 2013. Streamflow Prediction in Ungauged Basins: Review of Regionalization Methods. *Journal of Hydrologic Engineering*, 18(8):958-975.

Razavi, T. and P. Coulibaly, 2016. Improving streamflow estimation in ungauged basins using a multi-modelling approach. *Hydrological Sciences Journal*, 61(15):2668-2679.

Razavi, T. and P. Coulibaly, 2017. An evaluation of regionalization and watershed classification schemes for continuous daily streamflow prediction in ungauged watersheds. *Canadian Water Resources Journal*, 42(1):2-20.

Reed, S., V. Koren, M. Smith, Z. Zhang, F. Moreda, and D.J. Seo, 2004. Overall distributed model intercomparison project results. *Journal of Hydrology* 298(1-4):27-60.

Refsgaard, J.C., J.P. van der Sluijs, J. Brown, and P. van der Keur, 2006. A framework for dealing with uncertainty due to model structure error. *Advances in Water Resources* 29(11):1586-1597.

Rehn, A.C., R.D. Mazor, and P.R. Ode, 2015. The California Stream Condition Index (CSCI): A New Statewide Biological Scoring Tool for Assessing the Health of Freshwater Streams. California Water Boards Surface Water Ambient Monitoring Program (SWAMP) Technical Memorandum (TM) 2015-0002, 13 p., URL: https://www.waterboards.ca.gov/water_issues/programs/swamp/bioassessment/docs/csci_tech_memo.pdf

Rehn, A.C., P.R. Ode, J.T. May, 2005. Development of a Benthic Index of Biotic Integrity (B-IBI) for Wadeable Streams in Northern Coastal California and its Application to Regional 305(b) Assessment. California Water Boards Final Report. Surface Water Ambient Monitoring Program, California Environmental Protection Agency. URL: https://www.waterboards.ca.gov/water_issues/programs/swamp/docs/reports/final_north_calif_ibi.pdf.

Richter, B.D., J.V. Baumgartner, R. Wigington, and D.P. Braun, 1997. How much water does a river need? *Freshwater Biology* 37:231- 249.

Roussel, M.C., D.B. Thompson, X. Fang, T.G. Cleveland, and C.A. Garcia, 2005. Time-Parameter Estimation for Applicable Texas Watersheds. Texas Department of Transportation Research Report 0-4696-2.

Roy, A.H., M.C. Freeman, B.J. Freeman, S.J. Wenger, W.E. Ensign, and J.L. Meyer, 2005. Investigating hydrologic alteration as a mechanism of fish assemblage shifts in urbanizing streams. *Journal of the North American Benthological Association* 24(3):656-678.

Samuel., J., P. Coulibaly, and R.A. Metcalfe, 2011. Estimation of Continuous Streamflow in Ontario Ungauged Basins: Comparison of Regionalization Methods. *Journal of Hydrologic Engineering* 16(5):447-459.

Sanborn, S.C. and B.P. Bledsoe, 2006. Predicting streamflow regime metrics for ungauged streams in Colorado, Washington, and Oregon. *Journal of Hydrology* 325:241-261.

Santhi, C., P.M. Allen, R.S. Muttiah, J.G. Arnold, and P. Tuppad, 2008. Regional estimation of base flow for the conterminous United States by hydrologic landscape region. *Journal of Hydrology* 351(1-2):139-153.

Schueler, T., 2000. The Importance of Imperviousness. *Watershed Protection Techniques* 1(3): 100-111.

Sefick, S.A., L. Kalin, E. Kosnicki, B.P. Schneid, M.S. Jarrell, C.J. Anderson, M.H. Paller, and J.W. Feminella, 2015. Empirical Estimations of Stream Discharge using Channel Geometry in Low-Gradient, Sand-Bed Streams of the Southeastern Plains. *Journal of the American Water Resources Association* 51(4):1060-1071.

Sengupta, A., S.K. Adams, B.P. Bledsoe, E.D. Stein, K.S. McCune, and R.D. Mazor, 2018. Tools for managing hydrologic alteration on a regional scale: Estimating changes in flow characteristics at ungauged sites. *Freshwater Biology* 63(8) Special Issue: Evaluating and Managing Environmental Water Regimes in a Water-Scarce and Uncertain Future:769-785.

Sitterson, J., C. Knightes, R. Parmar, K. Wolfe, M. Muche, and B., Avant, 2017. An Overview of Rainfall-Runoff Model Types. U.S. Environmental Protection Agency (EPA)/600/R-17/482, Washington DC, 30p., URL: https://cfpub.epa.gov/si/si_public_record_report.cfm?dirEntryId=339328&Lab=NERL.

Skaugen, T. and D. Lawrence, 2017. The linear reservoir model: conceptual or physically based? 19th EGU General Assembly, proceedings from the conference held April 23-28, 2017 in Vienna, Austria. p. 7408.

Snelder, T.H., B.J.F. Biggs, and R.A. Woods, 2005. Improved eco-hydrological classification of rivers. *River Research and Applications* 21(6):609-628.

Snelder, T.H. and D.J. Booker, 2013. Natural Flow Regime Classifications are Sensitive to Definition Procedures. *River Research and Applications* 29(7):822-838.

Solans, M.A. and D. García de Jalón, 2016. Basic tools for setting environmental flows at the regional scale: application of the ELOHA framework in a Mediterranean river basin. *Ecohydrology* 9(8): 1517-1538.

State of California, 2019. California Water Plan Update 2018: Managing Water Resources for Sustainability. The Natural Resources Agency, Department of Water Resources, 80 p., URL: https://water.ca.gov/-/media/DWR-Website/Web-Pages/Programs/California-Water-Plan/Docs/Update2018/Final/California-Water-Plan-Update-2018.pdf.

State of Colorado, 2016. Water Supply Model (StateMod) Version 15. Department of Natural Resources, 323 p., URL: http://cwcbweblink.state.co.us/weblink/0/doc/198370/Electronic.aspx?searchid=ae150caf-e12e-4646-9e13-8b82d7b81ea4.

Stein, E.D., F. Federico, D.B. Booth, B.P. Bledsoe, C. Bowles, Z. Rubin, G.M. Kondolf, and A. Sengupta, 2012. Hydromodification Assessment and Management in California. Southern California Coastal Water Research Project (SCCWRP) Technical Report 667, 138 p., URL: ftp://ftp.sccwrp.org/pub/download/DOCUMENTS/TechnicalReports/667_CA_HydromodMgmt.pdf.

Stein, E.D., R.D. Mazor, A. Sengupta, K. McCune, B.P. Bledsoe, S.K. Adams, S.R. Eberhart, M.I. Pyne, P.R. Ode, and A.C. Rehn, 2017a. Development of Recommended Flow Targets to Support Biological Integrity Based on Regional Flow-ecology Relationships for Benthic Macroinvertebrates in Southern California Streams. Southern California Coastal Water Research Project (SCCWRP) Technical Report 974, 117 p., URL: http://ftp.sccwrp.org/pub/download/DOCUMENTS/TechnicalReports/974_RecomFlowTargetsForBioIntegrity.pdf.

Stein, E.D., A. Sengupta, R.D. Mazor, K. McCune, B.P. Bledsoe, and S.K Adams, 2017b. Application of regional flow-ecology relationships to inform watershed management decisions: Application of the ELOHA framework in the San Diego River watershed, California, USA. *Ecohydrology* 10(7):e1869

Tasker, G.D., 1982. Comparing Methods of Hydrologic Regionalization. *Water Resources Bulletin* 18(6):965-970.

Thiemann, M., M. Trosset, H. Gupta, and S. Sorooshian, 2001. Bayesian recursive parameter estimation for hydrologic models. *Water Resources Research* 37(10):2521-2535.

Tokar, A.S. and M. Markus, 2000. Precipitation-Runoff Modeling using Artificial Neural Networks and Conceptual Models. *Journal of Hydrologic Engineering* 5(2):156-161.

United States Army Corps of Engineers (USACE), 2000. Hydrologic Modeling System HEC-HMS Technical Reference Manual (CPD-74B). Institute for Water Resources Hydrologic Engineering Center. Davis, CA.

United States Department of Agriculture (USDA) Natural Resources Conservation Service (NRCS), 2007. Hydrologic Soil Groups. National Engineering Handbook, title 210—VI. Part 630 Hydrology, Chapter 7. Washington, DC. URL: https://directives.sc.egov.usda.gov/OpenNonWebContent.aspx?content=17757.wba.

USDA, 1986. TR-55: Urban Hydrology for Small Watersheds. Natural Resources Conservation Service., 164 p., URL: https://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/stelprdb1044171.pdf.

United States Census Bureau, 2012. California: 2010 Population and Housing Unit Counts. 2010 Census of Population and Housing, 189 p., URL: https://www.census.gov/prod/cen2010/cph-2-6.pdf.

United States Geological Survey (USGS), 2019a. The National Map. U.S. Department of the Interior, U.S. Geological Survey, URL: https://nationalmap.gov/.

USGS, 2019b. National Water Information System: Web Interface. U.S. Department of the Interior, U.S. Geological Survey, URL: https://waterdata.usgs.gov/nwis.

USGS, 2019c. StreamStats. U.S. Department of the Interior, U.S. Geological Survey, URL: https://streamstats.usgs.gov/ss/.

Van Dolah, R.F., J.L. Hyland, A.F. Holland, J.S. Rosen, and T.R. Snoots, 1999. A benthic index of biological integrity for assessing habitat quality in estuaries of the southeastern USA. *Marine Environmental Research* 48(4-5):269-283.

Vaze, J., P. Jordan, R. Beecham, A. Frost, and G. Summerell, 2011. Guidelines for Rainfall-Runoff Modelling: Towards Best Practice Model Application. eWater Cooperative Research Center, 50p., URL: https://ewater.org.au/uploads/files/eWater-Guidelines-RRM-(v1_0-Interim-Dec-2011).pdf

Venables, W.N. and B.D. Ripley, 2002. Modern Applied Statistics with S. Fourth Edition. Springer, New York.

Viney, N.R., 2013. Runoff Prediction in Support of the National Water Audit, Australia. In G. Blöschl, M. Sivapalan, T. Wagener, A. Viglone, and H. Savenije (Eds.). *Runoff Prediction in Ungauged Basins: Synthesis across Processes, Places, and Scales*, 78-80. Cambridge University Press, London, United Kingdom.

Vis, M., R. Knight, S. Pool, W. Wolfe, and J. Seibert, 2015. Model Calibration Criteria for Estimating Ecological Flow Characteristics. *Water* 7(5):2358-2381.

Vrugt, J.A., C.J.F. ter Braak, M.P. Clark, J.M. Hyman, and B.A. Robinson, 2008. Treatment of input uncertainty in hydrology modeling: Doing hydrology backward with Markov chain Monte Carlo simulation. *Water Resources Research* 44(12):W00B09.

Waananen, A.O., and J.R. Crippen, 1977. Magnitude and frequency of floods in California. Water-Resources Investigations Report 77-21, U.S. Department of the Interior, U.S. Geological Survey, 96 p., URL: https://pubs.er.usgs.gov/publication/wri7721.

Wagener, T. and A. Montanari, 2011. Convergence of approaches toward reducing uncertainty in predictions in ungauged basins. *Water Resources Research* 47(6).

Wagener, T., D.P. Boyle, M.J. Lees, H.S. Wheater, H.V. Gupta, and S. Sorooshian, 2001. A framework for development and application of hydrological models. *Hydrology and Earth System Sciences* 5(1):13-26.

Walsh, C.J., A.H. Roy, J.W. Feminella, P.D. Cottingham, P.M. Groffman, and R.P. Morgan, 2005. The urban stream syndrome: current knowledge and the search for a cure. *Journal of the North American Benthological Society* 24(3):706-723.

Walsh, C.J., A.K. Sharpe, P.F. Breen, and J.A. Sonneman, 2001. Effects of urbanization on streams of the Melbourne region, Victoria, Australia. I. Benthic macroinvertebrate communities. *Freshwater Biology* 46(4):535-551.

Weisberg, S.B., J.A. Ranasinghe, D.M. Dauer, L.C. Schaffner, R.J. Diaz, J.B. Frithsen, 1997. An estuarine benthic index of biotic integrity (B-IBI) for Chesapeake Bay. *Estuaries and Coasts* 20(1):149-158.

Western Regional Climate Tracker (WRCC), 2015. California Climate Tracker. Desert Research Institute, URL: http://www.wrcc.dri.edu/monitor/cal-mon/frames_version.html.

White, M.D. and K.A. Greer, 2006. The effects of watershed urbanization on the stream hydrology and riparian vegetation of Los Peñasquitos Creek, California. *Landscape and Urban Planning* 74(2):125-138.

Wilding, T.K., J.S. Sanderson, D.M. Merritt, S.B. Rood, and N.L. Poff, 2014. Riparian responses to reduced flood flows: comparing and contrasting narrowleaf and broadleaf cottonwoods. *Hydrological Sciences Journal* 59(3-4):605-617.

Wilson, M.A., and S.R. Carpenter, 1999. Economic Valuation of Freshwater Ecosystem Services in the United States: 1971-1997. *Ecological Applications* 9(3):772-783.

Wolch, J., 2007. Green Urban Worlds. *Annals of the Association of American Geographers* 97(2):373-384.

Wortman, R.T., 2005. Cross-Validation Error Statistic as a Tool for Model Evaluation. US Army Corps of Engineers Report, 7p.

Yankov, D., D. DeCoste, and E. Keogh, 2006. Ensembles of Nearest Neighbor Forecasts. 17th European Conference on Machine Learning, proceedings from the conference held September 18-22, 2006 in Berlin, Germany. p. 545-556.

Yarnell, S.M., E.D. Stein, J.A. Webb, T. Grantham, R.A. Lusardi, J. Zimmerman, R.A. Peek, B.A. Lane, J. Howard, and S. Sandoval-Solid, 2020. A functional flows approach to selecting ecologically relevant flow metrics for environmental flow applications. *River Research and Applications* 3(2):318-324.

Zhang, Y., A.H. Arthington, S.E. Bunn, S. Mackay, J. Xia, and M. Kennard, 2012. Classification of Flow Regimes for Environmental Flow Assessment in Regulated Rivers: The Huai River Basin, China. *River Research and Applications* 28(7):989-1005.

Zhang, Y.Q. and F.H.S. Chiew, 2009. Relative merits of different methods for runoff predictions in ungauged catchments. *Water Resources Research* 43(7):1-12.

Zimmerman, J.K.H., D.M. Carlisle, J.T. May, K.R. Klausmeyer, T.E. Grantham, L.R. Brown, and J.K. Howard, 2017. Patterns and magnitudes of flow alteration in California, USA. *Freshwater Biology* 63(8):859-873.

Zimmerman, J.K.H., B.H. Letcher, K.H. Nslow, K.A. Lutz, and F.J. Magilligan, 2010. Determining the Effects of Dams on Subdaily Variation in River Flows at a Whole-Basin Scale. *River Research and Applications* 26:1246-1260.

Zorn, T.G., P.W. Seelbach, E.S. Rutherford, T.C. Willis, S.T. Cheng, and M.J. Wiley, 2008. A Regional-scale Habitat Suitability Model to Assess the Effects of Flow Reduction on Fish Assemblages in Michigan Streams. Michigan Department of Natural Resources Fisheries Division Fisheries Research Report 2089, 50p., URL: https://www.michigan.gov/documents/dnr/RR2089_268570_7.pdf.

Zvolensky, M., S. Kohnova, K. Hlavcova, J. Szolgay, and J. Parajka, 2008. Regionalisation of rainfall-runoff model parameters based on geographical location of gauged catchments. *Journal of Hydrology and Hydromechanics* 56(3):176-189.

Appendix A

A.1 Additional Background

Stream classification is typically the second step of ELOHA, following a hydrologic foundation but prior to flow alteration and flow-ecology linkages. Streams can be classified at any scale (Poff *et al.*, 2010) ranging from global (Puckridge *et al.,* 1998), to national (Poff, 1996), across multiple states (McManamay *et al.*, 2012), an individual state (Liermann *et al*., 2012), or a singular river basin (Belmar *et al.*, 2011). The Pyne *et al.* (2017) classification facilitates estimating flow alteration at minimally disturbed gage sties; however, ungaged basins are not considered.

ELOHA requires a regional network of paired bioassessment and streamflow data, while regionalization of ungaged basins utilizes a network of hydrologic models. Different types of statistically-based regionalization have been used within stream classification to model streamflow metrics in ungaged basins (Brown *et al.*, 2014; Carlisle *et al.*, 2010a; Lane *et al.*, 2017; Santhi *et al.*, 2008), but has not been used to model entire time series of discharge with rainfall-runoff models.

A.2 Additional Results

A.2.1 Combined Inductive and Deductive Approaches

An expanded cluster analysis forcing drainage area, HGC, RBI, and < 1 cfs as predictors did not improve classification over individual inductive and deductive approaches (Figure A2). The combined cluster analysis was dominated by drainage area and produced the same clusters as the traditional deductive approach. Using RBI and < 1 cfs in multinomial logistic regression to predict deductively produced clusters of watershed characteristics was unsuccessful with prediction errors between 52% and 80% (between 5 and 12 sites correctly predicted). Similarly, drainage area and HGC were unable to predict inductively produced clusters of flow metrics with prediction errors between 52% and 60% (between 10 and 12 sites correctly predicted). A new multinomial regression model was able to predict flow metric clusters with 16% error (21 sites correctly predicted) using drainage area, % clay soil, minimum elevation, and annual minimum precipitation (Figure A3). Drainage area was the strongest predictor for discerning flow permanence with intermittent streams containing the smallest basins, perennial having the biggest, and ephemeral in between. Intermittent streams were also generally found at higher elevations where soils contain less clay and annual minimum precipitation volumes are generally larger (Figure A3). A second successful multinomial regression model containing drainage area, % silt soil, baseflow index, and relative humidity was able to predict whether a gage was classified as reference by GAGES-II with 12% error (22 sites correctly predicted) (Figure A4). No combination of RBI, < 1 cfs, drainage area, and HGC could predict a gage's reference status, wherein all sites were classified as non-reference in multinomial logistic regression, with some minor exceptions. Drainage area provided the clearest distinction of reference status, with smaller basins tending to be less altered than larger ones. These smaller basins contain higher silt content and Baseflow Indices. Flashiness and flow permanence metrics were not different between reference and non-reference gages (Figure A4). The reference class provided a significantly smaller ACE (0.4, Figure A4) than the non-reference class (1.9, Figure A4). The ACE across both sites is much closer to the non-reference class (1.4, Figure A4).

A.3 Additional Discussion

A.3.1 Stream classification in coastal southern California

The classifications presented in this study each provide important complementary information characterizing streams in coastal southern CA. Three distinct classes of flow permanence (ephemeral, intermittent, and perennial) from the inductive approach provided the strongest individual classification, highlighting the heterogeneity of flow permanence in So. CA and its importance in management. Results showing inductive classification outperforming deductive agree with a study across New Zealand (Snelder and Booker, 2013) and support the idea put forth in ELOHA that stream classification should primarily be focused on hydrologic similarity, when feasible (Poff *et al.*, 2010). This study has shown that deductive classification with watershed area and percentage of relatively low infiltration high runoff soils, as HGC (USDA NRCS, 2007), can provide a reliable alternative classification to that of ecologically-relevant flow metrics. Drainage area is an important watershed characteristic that appeared in all classifications utilizing watershed data. It has a complex spatial relationship with sub-regional topography and precipitation wherein bigger basins typically encompass lower, flatter elevations with less orographic influence and associated rainfall. Despite receiving less average rainfall, streams in these large watersheds typically have greater flow magnitude and permanence due to the slow accumulation of baseflow from large contributing areas. Soil metrics, such as HGC, play a role in this phenomenon by controlling how rainfall is partitioned into fast-acting overland flow or gradual baseflow. Many quantifiable watershed characteristics are fundamental to the hydrologic processes that control how rainfall becomes streamflow; however, drainage basin size and soil type certainly play important roles in controlling the magnitude, duration, frequency, timing, and rate of change of streamflow, and are shown to help accurately separate streams in So. CA into distinct classes as potential management units. Deductive classification is performed less frequently than inductive approaches and does not always characterize hydrologic similarity (Carlisle *et al.*, 2010a; Snelder *et al.*, 2005), but it still provides important information regarding spatial variation of streamflow (Olden *et al.*, 2012), specifically in reference to drainage area and HGC in So. CA. While comparison of different stream classification approaches for the same sites has not been frequently performed in the literature, results from this study agree with the findings of Snelder and Booker (2013) wherein a variety of diverse classifications, including inductive and deductive approaches, can be accurately implemented across a region.

Inductive and deductive approaches to stream classification each have their merits and drawbacks (Olden *et al.*, 2012), but joining them into one combined classification can provide synergistic information regarding regional streamflow. The two successful traditional combined classifications produced for So. CA in this study (Figure A3 and Figure A4) are similar to the state-wide classifications produced by Lane *et al.* (2017) and Pyne *et al.* (2017), respectively. Like this study, Lane *et al.* (2017) used watershed characteristics to predict hydrologic classes generated by the k-means algorithm; however, they considered a suite of hydrologic metrics more relevant to the entire state, not just streamflow flashiness and permanence. This led them to produce a wide range of hydrologic regimes characterizing typical annual hydrographs. These general flow regimes are useful at a state-wide scale but provide no information regarding the heterogeneity of the south coast. According to Lane *et al.* (2017), the hydrologic regime for practically every stream in the coastal southern CA can be characterized as "rain and seasonal groundwater", which occur at low coastal elevations with limited winter precipitation. These streams are abundant in the south coast, but streams with higher permanence also exist throughout the region (Devil Canyon, San Luis Rey, Sandia, and Santa Margarita Sump) and are more characteristic of Lane *et al.* (2017)'s "perennial groundwater and rain" regime containing more stable flows and little clay soil at low elevations. Still other streams in the south coast resemble the "flashy, ephemeral rain" hydrologic regime produced by Lane *et al.* (2017), including Carpinteria and Santiago, with its characteristic low elevation and high clay soil content. Some of the watershed characteristics used to predict flow permanence in this study were the same as those used to predict general hydrologic regime across the entire state. Drainage area, percent clay soil, basin elevation, and a measure of precipitation were common predictors of hydrologic regime between studies (Lane *et al.*, 2017). This consistency across two different, but overlapping, geographic scales suggests the potential for a new type of nested classification, wherein similar watershed characteristics are used to differentiate general large-scale hydrologic regimes, and small-scale regimes specific to local management needs. This would help stratify stream management for different stakeholders who might consider management at different scales.

Essentially the inverse of Lane *et al.* (2017), Pyne *et al.* (2017) classified all California stream segments by watershed characteristics, and then predicted class membership with hydrologic metrics.

This process by Pyne *et al.* (2017) was inaccurately duplicated in So. CA using two flow metrics known to be ecologically-relevant in the south coast (Gasith and Resh, 1999; Mazor *et al*., 2018; Parker *et al*., 2019); however, the traditional deductive classification produced in this study (Figure 2.4) had some similarities with the state-wide classification. Winter precipitation, geology, soil content, and mean elevation separated stream segments across the state (Pyne *et al.*, 2017), while basin area and soil content provided distinct classes in the south coast. Furthermore, the multinomial logistic regression model predicting GAGES-II reference status in the south coast (Figure A4) provides information complementary to the Random Forest (RF) model produced by Pyne *et al.* (2017), wherein they considered a new set of watershed characteristics describing disturbance, and found the total volume of reservoirs, impervious cover, freshwater withdrawals, high-intensity developed land use, cropland, and population density were most significant at determining reference status for all stream segments across CA. Estimating flow alteration and its causes was outside the scope of this study; however, a logistic regression model containing typical watershed characteristics was able to accurately determine USGS streamflow gages' reference status. The similarities between deductive classification and a combined approach for determining reference status in Pyne *et al.* (2017) and this study, while not as strong as similarities between inductive and combined classification in Lane *et al.* (2017), again suggest the potential behind a type of nested classification. It is not surprising that inductively-based combined classification, as was performed in Lane *et al.* (2017) and Figure A3, provides more similar results than deductively-based combined classification between state-wide and south coast regional scales. California is an extremely heterogeneous state, but the diversity of streamflow in the south coast is more representative of state-wide streamflow than the range of watershed characteristics is representative of the entire state. This finding provides more support behind primarily classifying streams using hydrologic similarity, especially in heterogeneous landscapes when different geographic scales may be important.

A comparison between traditional combined classification approaches (Figure A2, Figure A3, and Figure A4) and Hydrologic Model-based Classification (Figure 2.5) shows similar and complementary information regarding south coast streams. Drainage area and soil content were the most consistent watershed characteristics used to classify streams from the region, regardless of approach. In addition to these important metrics, HMC classes were predicted by two measures of precipitation: annual minimum and annual average. Precipitation is the primary input to rainfall-runoff models, and so its importance for determining low-error classes is not surprising. A deeper look into the average values of watershed characteristics for each class shows wider spread and stronger cluster identity for HMC. Multinomial logistic regression provided clearer separation of HMC classes (Figure 2.5) than flow permanence (Figure A3) and reference status (Figure A4) classes, resulting in more accurate class prediction with an error of 4% compared to 16% and 12%, respectively.

A.3.2 Stream classification for regionalizing ungaged basins

Because the expanded cluster analysis forcing drainage area, HGC, RBI, and < 1 cfs as predictors produced the same classification as the deductive approach, it also contained an ACE of 0.4 (Figure A2). Similarly, combined classification predicting streamflow classes using multinomial logistic regression produced the same ACE as traditional inductive classification (ACE 0.6; Figure A3).

While Hydrologic Model-based Classification demonstrated the most overall potential in ungaged basins, it did not produce all the lowest error classes. The 0.5 ACE produced by HMC class five (Figure 2.5) was larger than a few traditional classes including the inductive intermittent class (ACE 0.2; Figure 2.3), the deductive small basin/low HGC and large basin/high HGC classes (ACE 0.3 and 0.2, respectively; Figure 2.4), and the combined classification reference class (ACE 0.4; Figure A4). Furthermore, deductive medium-sized basins/high HGC and small basins/high HGC produced the same ACE values of 0.5 (Figure 2.4). HMC class five contains a comparably larger ACE because it has many more sites than any other class, regardless of classification. The most sites belonging to one of the six traditional classes with an ACE of 0.5 or below is 7, about half the 13 sites in hydrologic model-based class five. While ACE is normalized to the number of sites in a class, the diversity of a class increases with the number of sites because no two sites are alike, especially in the south coast. The increased heterogeneity associated with larger classes inherently reduces their predictive ability. The presence of one large class is unique to HMC and reflects the tremendous accuracy of the other four classes (ACE 0.0 to 0.1, Figure 2.5), which more than compensate for class five to overall produce the most accurate classification for estimating flow in ungaged basins.

A.4 Additional Tables and Figures

Table A1: Study sites details. Drainage area data are from the National Hydrography Dataset Plus Version 2 (NHDPlus V2) (McKay et al., 2012), impervious data are from StreamStats (USGS, 2019c), and all other watershed characteristics, including reference gage status, are from GAGES-II (Falcone, 2011). Mean Annual Flowrates were computed from gage data.

Site Name	**USGS Gage**	**Ref-erence**	**Area (km^2)**	**Imper vious %**	**HGC %**	**Clay %**	**Silt %**	**Sand %**	**Basin Min Elevation (m)**	**Catchment Annual Min Precip (cm)**	**Catchment Annual Avg Precip (cm)**	**BFI %**	**RH %**	**Mean Annual Flowrate (cfs)**		
														WY 2005	**WY 2006**	**WY 2007**
Andreas	10259000	Ref	23.2	0.0	12	10	30	59	177	1.0	32	49	42	6.1	2.2	0.9
Arroyo Seco	11098000	Ref	42.5	0.5	43	18	48	34	398	1.0	63	38	54	52	8.6	0.9
Arroyo Trabuco	11047300	Non-ref	141.4	19.9	16	25	43	33	18	1.0	34	24	62	69	13	5.1
Campo	11012500	Non-ref	222.3	7.0	7.4	10	20	70	623	2.0	42	41	47	2.0	0.4	0.1
Carpinteria	11119500	Non-ref	45.4	0.1	23	23	42	34	6.6	0.0	44	30	54	18	3.7	0.0
Deep Creek	10260500	Non-ref	354.0	2.4	52	12	24	64	915	2.0	23	33	40	171	63	7.8
Devil Canyon	11063680	Non-ref	14.7	0.7	49	14	32	54	598	3.0	80	40	47	2.7	4.9	2.1
East Twin	11058500	Non-ref	23.1	0.7	41	14	31	56	483	2.2	61	34	45	2.7	5.4	1.6
Jamul	11014000	Non-ref	182.9	0.5	33	19	39	43	153	2.0	36	34	56	23	0.1	0.0
Lytle	11062000	Non-ref	119.8	0.4	16	11	31	58	725	2.0	91	54	48	37	32	3.1
Matilija	11114495	Ref	128.5	0.0	37	22	42	36	348	0.0	72	39	50	156	37	4.3
Plunge	11055500	Non-ref	44.2	1.3	40	13	31	55	485	2.4	45	29	43	14	7.9	2.1
Poway	11023340	Non-ref	110.0	21.8	18	23	38	39	75	1.0	34	24	60	36	7.3	4.5
Rainbow	11044250	Non-ref	27.0	4.3	29	15	31	54	161	1.0	45	30	59	16	1.4	0.4
San Luis Rey	11042000	Non-ref	1433.8	3.1	32	15	28	57	1.2	1.0	31	35	53	229	29	9.7
San Mateo	11046300	Ref	210.2	0.1	50	17	35	47	120	1.7	46	26	56	90	3.3	0.1
Sandia	11044350	Non-ref	51.1	1.3	37	19	39	42	124	1.0	45	28	59	30	6.1	4.0
San Jose	11120500	Ref	15.8	0.4	9.5	20	45	35	26	0.7	48	19	61	14	3.0	0.2

Santa Margarita Sump	11044300	Non-ref	1576.9	4.6	29	15	30	55	99	1.0	43	29	51	82	18	9.4
Santa Ysabel	11025500	Non-ref	290.9	0.1	52	16	31	52	247	2.0	43	32	49	29	1.7	0.0
Santa Maria	11028500	Non-ref	147.7	2.6	38	17	32	51	397	2.0	44	27	53	16	0.4	0.1
Santiago	11075800	Non-ref	32.9	0.1	7.8	23	46	31	374	1.0	49	25	59	20	1.8	0.0
Sespe Fillmore	11113000	Non-ref	651.0	0.1	41	22	42	36	168	0.0	52	45	45	515	211	15
Sespe Wheeler Springs	11111500	Ref	131.9	0.1	48	22	41	37	1028	0.0	69	40	45	87	22	1.4
Sweetwater Descanso	11015000	Ref	126.0	0.3	26	13	25	62	754	3.0	62	40	46	21	3.7	1.0

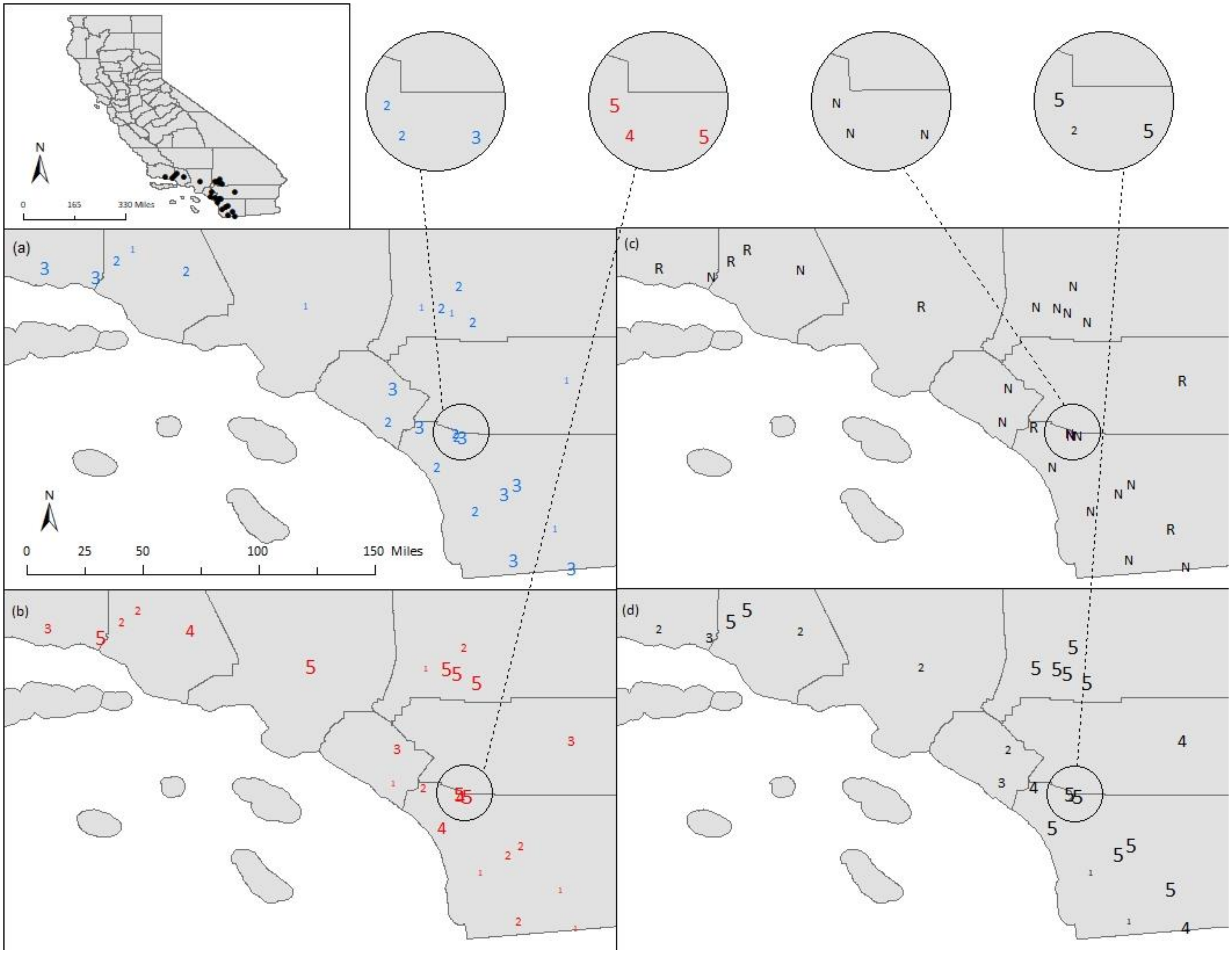

Figure A1: Geographical distribution of classes, specifically for (a) traditional inductive approach; (b) traditional deductive approach; (c) GAGES-II reference sites; (d) new hydrologic model-based method.

Class	Drainage Area (km^2)	HGC %	RBI	< 1 cfs	Avg Cluster Error
1	143.9	17	0.09	0.48	0.6
2	206.6	44	0.08	0.006	0.5
3	24.0	10	0.09	0.55	0.3
4	1220.6	34	0.10	0.33	0.2
5	35.4	37	0.09	0.28	0.5
All					**0.4**

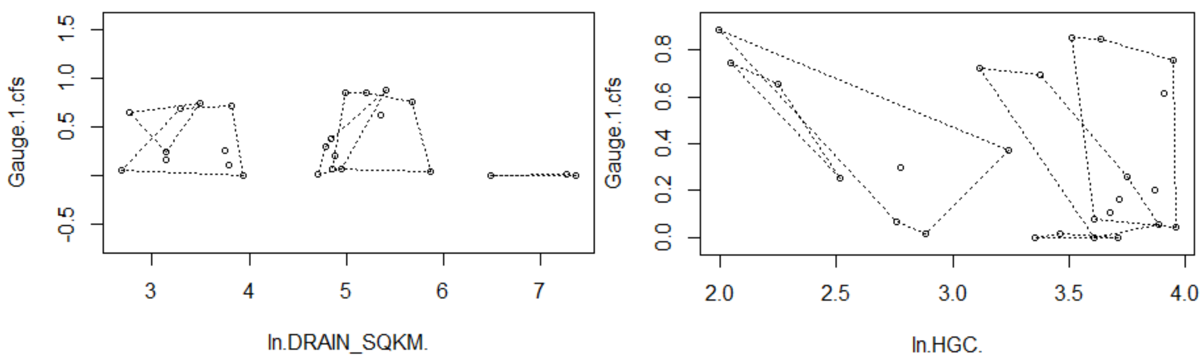

Figure A2: Results for expanded clustering as a combined inductive and deductive approach to traditional classification

Class	Drainage Area (km^2)	Clay %	Min Elevation (m)	Annual Min Precip (cm)	Avg Cluster Error
1	77.8	15	594	1.5	0.2
2	450.5	18	283	1.2	0.9
3	130.6	18	234	1.4	0.6
All					**0.6**

Logistic Regression Landscape Variable	Definition	Source
DRAIN_SQKM	Total upstream drainage area (km^2)	NHDPlus V2 (McKay *et al.*, 2012)
CLAYAVE	Percentage of clay soil (%)	GAGES-II (Falcone, 2011)
ELEV_MIN_M_BASIN	Minimum basin elevation (m)	GAGES-II (Falcone, 2011)
CAT_AnnMinPrecip	Mean annual minimum precipitation of NHD catchment (cm)	GAGES-II (Falcone, 2011)

Figure A3: Results for predicting inductively produced streamflow classes, as a combined inductive and deductive approach to traditional classification.

Class	Drainage Area (km^2)	Silt %	BFI %	RH %	RBI	< 1 cfs	Avg Cluster Error
Ref	96.9	38	36	51	0.08	0.35	0.4
Non-ref	303.8	34	33	52	0.10	0.35	1.9
All							**1.4**

Logistic Regression Landscape Variable	Definition	Source
DRAIN_SQKM	Total upstream drainage area (km^2)	NHDPlus V2 (McKay *et al.*, 2012)
SILTAVE	Percentage of clay soil (%)	GAGES-II (Falcone, 2011)
BFI_AVE	Baseflow Index (% as baseflow/total flow)	GAGES-II (Falcone, 2011)
RH_BASIN	Mean basin relative humidity (%)	GAGES-II (Falcone, 2011)

Figure A4: Results for predicting reference site status, as a combined inductive and deductive approach to traditional classification.

Table A2: Calibration results with model % errors of RBI and < 1 cfs relating model accuracy to gaged data. EFCC (%) refers to Equation 2.3. Sites with high flashiness, "H", contained a gaged RBI greater than 0.125 during WY2005-2007; sites with low flashiness, "L", had an RBI less than 0.075; sites with average flashiness, "A", had an RBI between 0.075 and 0.125.For flow permanence, ephemeral streams are represented by "E" and contained gaged streamflow < 1 cfs more than half the time during WY2005-2007; perennial streams, "P", had flow < 1 cfs less than 10% of time; and intermittent streams, "I", had streamflow < 1 cfs 10%-50% of time.

Site Name	Gage RBI	Flashiness	Model % Error RBI	Gage < 1 cfs	Flow Permanence	Model % Error < 1 cfs	EFCC (%)
Andreas	0.05	L	0.0	0.25	I	0.0	0.0
Arroyo Seco	0.06	L	0.3	0.26	I	0.1	0.2
Arroyo Trabuco	0.16	H	0.1	0.07	P	0.6	0.4
Campo	0.04	L	0.1	0.89	E	0.1	0.1
Carpinteria	0.08	A	0.1	0.72	E	0.0	0.1
Deep Creek	0.12	A	0.9	0.04	P	0.5	0.7
Devil Canyon	0.03	L	0.3	0.05	P	0.6	0.4
East Twin	0.08	A	0.4	0.16	I	0.2	0.3
Jamul	0.12	A	0.2	0.85	E	1.1	0.7
Lytle	0.05	L	0.3	0.30	I	0.1	0.2
Matilija	0.05	L	0.3	0.07	P	0.3	0.3
Plunge	0.08	A	0.4	0.10	I	0.0	0.2
Poway	0.19	H	0.6	0.02	P	1.8	1.2
Rainbow	0.20	H	0.6	0.69	E	0.2	0.4
San Luis Rey	0.04	L	0.2	0.02	P	0.8	0.5
San Mateo	0.07	L	0.1	0.62	E	0.0	0.1
Sandia	0.09	A	0.9	0.00	P	0.0	0.4
San Jose	0.16	H	0.1	0.66	E	0.0	0.0
Santa Margarita Sump	0.13	H	1.1	0.00	P	0.0	0.6
Santa Ysabel	0.11	A	0.5	0.76	E	0.0	0.3
Santa Maria	0.09	A	0.1	0.85	E	0.0	0.1
Santiago	0.07	L	0.3	0.75	E	0.0	0.2
Sespe Fillmore	0.09	A	0.3	0.00	P	0.0	0.1
Sespe Wheeler Springs	0.06	L	0.4	0.20	I	0.1	0.2
Sweetwater Descanso	0.07	L	0.6	0.37	I	0.9	0.7
Mean			**0.4**			**0.3**	**0.3**
Median			**0.3**			**0.1**	**0.3**

Appendix B

B.1 Additional Background

By definition, watershed models can never fully represent reality, but they have improved significantly since their laboratory origins as scaled-down physical representations of watershed processes (Chery Jr., 1966). Complex modern watershed modeling methods can be divided into three general groups: empirically-based statistical models, physically-based models conserving mass, momentum, and energy, and process-based conceptual models with simpler physics (Sitterson *et al.*, 2017). Conceptual and physically-based watershed models, beginning with the Stanford Watershed Model (Crawford and Linsley, 1966), are often referred to as rainfall-runoff models and use mathematic equations to simulate processes controlling streamflow, which grounds them in theoretical physics and enables them to produce entire time series of discharge from which any flow metric can be computed (Vaze *et al.*, 2011). Conversely, statistical models can typically only estimate specific flow metrics for which they have been conditioned (Helsel and Hirsch, 2002). Conceptual models offer greater flexibility than their physically-based counterparts by providing varying levels of physical complexity (Duan *et al.*, 1992), which becomes important when data are too scarce or unreliable for a full physically-based model.

The physical complexity of conceptual models can be separated into three modeling structures: lumped, distributed, and semi-distributed (Sitterson *et al.*, 2017). The simplest type, lumped, utilizes basin-scale data with basic physics. More complicated distributed structures are required for physically-based models and can be use in conceptual models. These structures implement high-resolution spatially-variable data across a user-generated grid (Sitterson *et al.*, 2017), and can produce more accurate results when operated by an experienced modeler (Carpenter and Georgakakos, 2006), but model complexity is not always correlated with accuracy (Beven, 1989; Grayson *et al*., 1992; Reed *et al.*, 2004). A semi-distributed conceptual watershed model aggregates multiple sub-basin lumped models to provide finer detail of a watershed than a singular lumped conceptual model, without the input and computational burdens of fully-distributed physically-based models (Sitterson *et al.*, 2017). Even though they are not fully physically-based, lumped and semi-distributed conceptual models are process-oriented and can still provide a physical understanding of the mechanisms that transform rainfall into streamflow, especially those related to losses (Vaze *et al.*, 2011) and baseflow (Skaugen and Lawrence, 2017).

Conceptual models further benefit from manual or automatic calibration to increase their accuracy through adjusting non-calculated parameter(s) such that modeled hydrographs better match observed data (Vaze *et al.*, 2011). These parameter(s) are used in the simplified conservation of mass, momentum, and energy equations, but cannot be calculated directly with available watershed data. Guided by measured streamflow, calibration of rainfall-runoff models optimizes some performance metric(s), typically a singular measure of overall hydrograph fit (Bardossy, 2007; Beven, 2012) such as the Nash-Sutcliffe Efficiency (NSE) (Jain and Sudheer, 2008; Nash and Sutcliffe, 1970). NSE has a long history as a commonly applied model assessment criterion, but it and other similar overall-fit metrics based on mean flow are biased towards high flow accuracy (Jain and Sudheer, 2008; Legates and McCabe, 1999). Prioritizing the accuracy of high flows is ideal for managing large quantities of water, such as for flood control, but is not preferred for many other model applications, such as managing environmental flows, wherein the accuracies of ecologically-relevant metrics, often describing low flows and streamflow variability, are most important (Parker *et al*., 2019). Flow regime elements most crucial to environmental flow studies are typically not specifically considered in best overall-fit metrics.

Unfortunately, gaged streamflow is not always available to calibrate rainfall-runoff models and so models must be used to make streamflow predictions in ungaged basins (PUB). Statistical, conceptual, and physically-based modeling techniques are used to make PUBs (Blöschl *et al.*, 2013;

Carlisle *et al.*, 2010a; Kennen *et al.*, 2008; Peel and Blöschl, 2011; Razavi and Coulibaly, 2013; Sanborn and Bledsoe, 2005; State of Colorado, 2016); however, conceptual models typically offer flexibility without the confounding effects of too many parameters. As such, they are more commonly used than statistical or physically-based models, respectively (Vaze *et al.*, 2011). Some of this flexibility involves easily changing parameters to estimate historical or future conditions, adjusting time scales, producing time series data, and assessing management scenarios (Blöschl *et al.*, 2013; Buchanan *et al.,* 2013; Hydrologic Engineering Center, 2015; Kendy *et al.*, 2012; Pomeroy *et al.*, 2008; Razavi and Coulibaly, 2013; State of Colorado, 2016). To this point, a calibrated conceptual model outperformed empirical statistical models in estimating ungaged streamflow across New Zealand (Booker and Woods, 2014).

The physically-based elements of conceptual models, which allow them to accurately integrate landscape data into parameter estimates, reduce parameter uncertainty compared to statistical models during regionalization (Vaze *et al.*, 2011). Physically-based models have the potential for the same benefits during regionalization, but usually contain too many parameters for sensical donation to a model of an ungaged basin (Vaze *et al.*, 2011).

To quantify and reduce parameter uncertainty, Bayesian statistics are commonly applied to iteratively update parameter values based on the probability of their accuracy, given some prior knowledge of their values (Gelman *et al.*, 2013). The Generalized Likelihood Uncertainty Estimation (GLUE) methodology (Beven and Binley, 1992) and the Bayesian Recursive Estimation (BaRE) framework (Thiemann *et al*. 2001) estimate parameter uncertainty using Bayesian statistics, and can be used to reduce modeled flow inaccuracies caused by parameter uncertainty. These two approaches for quantifying and reducing uncertainty are typically applied to individual models and not across a region. Knowledge of parameters from other models within a region may be used to inform a prior likelihood of estimated parameter accuracy, but GLUE and BaRE are not designed explicitly for regionalization. Furthermore, as with other methods for quantifying uncertainty, such as the Metropolis algorithm

(Kuczera and Parent, 1998) or the differential evolution adaptive Metropolis (DREAM) (Vrugt *et al.*, 2008), GLUE and BaRE rely on high-dimensional automated Monte Carlo sampling to assess the likelihood a certain parameter or parameter set fits a model. This approach for quantifying uncertainty is computationally intensive, especially in heterogeneous regions where a modeler might have a poor understanding of prior parameter likelihoods and is not incorporated well into many rainfall-runoff modeling software.

B.2 Additional Tables and Figures

Logistic Regression Landscape Variable	**Definition**	**Source**
DRAIN_SQKM	Total upstream drainage area (km^2)	NHDPlus V2 (McKay *et al.*, 2012)
SANDAVE	Percentage of sandy soil (%)	GAGES-II (Falcone, 2011)
PPTAVG_CAT	Mean annual precipitation of NHD catchment (cm)	GAGES-II (Falcone, 2011)
CAT_AnnMinPrecip	Mean annual minimum precipitation of NHD catchment (cm)	GAGES-II (Falcone, 2011)

Cluster	**Drainage Area (km^2)**	**Sand %**	**Annual Avg Precip (cm)**	**Annual Min Precip (cm)**
1	146.4	41	35	1.9
2	463.8	38	51	1.0
3	93.4	33	39	0.6
4	151.9	59	40	1.6
5	222.5	52	55	1.8

Table B1: HMC results for creating ensembles with landscape variables to place ungaged sites into reciprocally low error clusters.

Appendix C

C.1 Additional Background

The ELOHA framework provides a consensus methodology for developing environmental flow standards across an entire region. The framework can be broken down into four major components: 1) Hydrologic foundation; 2) River classification; 3) Flow alteration; 4) Flow-ecology linkages (Poff *et al*., 2010).

Flow-ecology relationships provide important management insight, grounded in a scientific approach, to connections between streamflow and ecological integrity, for modified and natural rivers (Poff *et al*., 2010). Flow metrics comprising all elements of the flow regime have been used to develop both quantitative and qualitative flow-ecology relationships with many types of bioassessment endpoints, including macroinvertebrates, fish, and vegetation (Poff and Zimmerman, 2010). When available, existing literature of regional flow-ecology relationships is recommended as a foundation for developing environmental flows (Davies *et al.*, 2014). To produce flow-ecology relationships, two pieces of data are needed for the same stream site: a measure of flow alteration and a measure of ecological change. With these two data for multiple stream sites in a region, relationships can be quantified with a variety of statistical techniques (Poff and Zimmerman, 2010). The most common approach uses correlation between flow alteration and ecological change (Booth *et al.*, 2004), sometimes in concert with a type of regression (Buchanan *et al.*, 2013; Kennen *et al.*, 2010; Konrad *et al.*, 2008; McManamay *et al.*, 2013; Pomeroy *et al.*, 2008).

Developing flow-ecology relationships begins with establishing bioassessment sites for measuring ecological change. Rapid Bioassessment Protocols (Barbour *et al.*, 1999) provide a framework of standardized methods for determining the degradation of ecological integrity from reference conditions by deriving ecological metrics from biological surveys. Ideally, ecological metrics calculated from biological surveys quantify not just taxonomy, but also trait-based functional attributes, which can provide more generalized flow-ecology relationships that are transferable across spatial and temporal scales (Poff, 1997). Nonetheless, results from taxonomic biological surveys may be comparable across regions when taxonomic completeness is considered, as measured by the proportion of expected taxa to observed taxa $\left(\frac{\text{Observed}}{\text{Expected}}, \frac{\text{O}}{\text{E}}\right)$ (Hawkins, 2006).

Competing statistical and rainfall-runoff modeling approaches have produced different estimates of the environmental flow regime within studies spanning the entirety of New Zealand (Booker and Woods, 2014) and the United States (Carlisle *et al.*, 2010a), Kentucky (Murphy *et al.*, 2013), and coastal southern CA (Ch. 2).

C.2 Additional Discussion

C.2.1 Flow-ecology relationships in ungaged basins

Creating flow-ecology relationships in ungaged basins is made easier with a multimetric bioassessment index. When normalized by benchmark ecological conditions, multimetric indices such as CSCI, reduce uncertainty regarding which ecological metrics and formulations of those metrics should be prioritized over others to create flow-ecology relationships. Without the multimetric index in this study, a full analysis of ecological alteration would complicate results significantly. Despite these advantages of multimetric indices over simple taxonomic abundance or richness, many previous environmental flow studies quantifying flow-ecology relationships have not considered metrics describing function attributes and habitat characterizations. This can be seen for cottonwood trees in Colorado (Wilding *et al.*, 2014), for fish in Michigan (Zorn *et al.*, 2008) and the Central Valley of California (Marchetti and Moyle, 2001), for fish and riparian cover in the Upper Tennessee River basin (McManamay *et al.*, 2013), and for both periphyton and invertebrates in New Zealand (Clausen and Biggs, 1997). To take advantage of more generalized flow-ecology relationships, some studies have incorporated flow or habitat guilds and life history strategies of fishes, such as in the SW US (Chen and Olden, 2017), North Carolina (Phelan *et al.*, 2017), Massachusetts (Armstrong *et al.*, 2011), Georgia (Freeman and Marcinek, 2006). Similarly, abundance and richness data of invertebrates belonging to Ephemeroptera, Plecoptera, and Trichoptera (EPT) taxa, and other characteristic functional, behavioral, or habitat groups, are often used as indicators of water quality (Barbour *et al.*, 1999), and have been used extensively in generalized flow-ecology studies, such as in the Piedmont (Pomeroy *et al.*, 2008) and entire state of North Carolina (Phelan *et al.*, 2017), across the Western US (Konrad *et al.*, 2008), within the Potomac River basin (Buchanan *et al.*, 2013), and across New Jersey (Kennen *et al.*, 2013; Kennen *et al.*, 2010). While these ecological metrics generally provide more robust relationships than metrics of individual taxonomy, they still require estimates of ecological condition for both baseline and altered flows. As such, multimetric indices eliminate the need for two distinct measures of ecological condition per site. Examples of multimetric indices include the Great-River Macroinvertebrate Indices of Condition (GRMINs) (Angradi *et al*., 2009), Canadian Ecological Flow Index (CEFI) (Armanini *et al.*, 2011), New Jersey Impairment Score (NJIS) (Kennen *et al.*, 2013), and Regional Biological/Invertebrate Indices of Biotic Integrity (IBI) spanning the US from the Tennessee Valley (Kerans and Karr, 1994) to the Chesapeake Bay (Weisberg *et al.*, 1997), to Northern (Rehn *et al.*, 2005) and Southern coastal California (Ode *et al.*, 2005), to the southeast (Van Dolah *et al.*, 1999), and all regions in between. The CSCI has been applied across California from the Sierra Nevada Mountains (Carlisle *et al.*, 2016) to So. CA (Mazor *et al.*, 2018; Stein *et al.*, 2017a; Stein *et al.*, 2017b).

C.3 Additional Tables and Figures

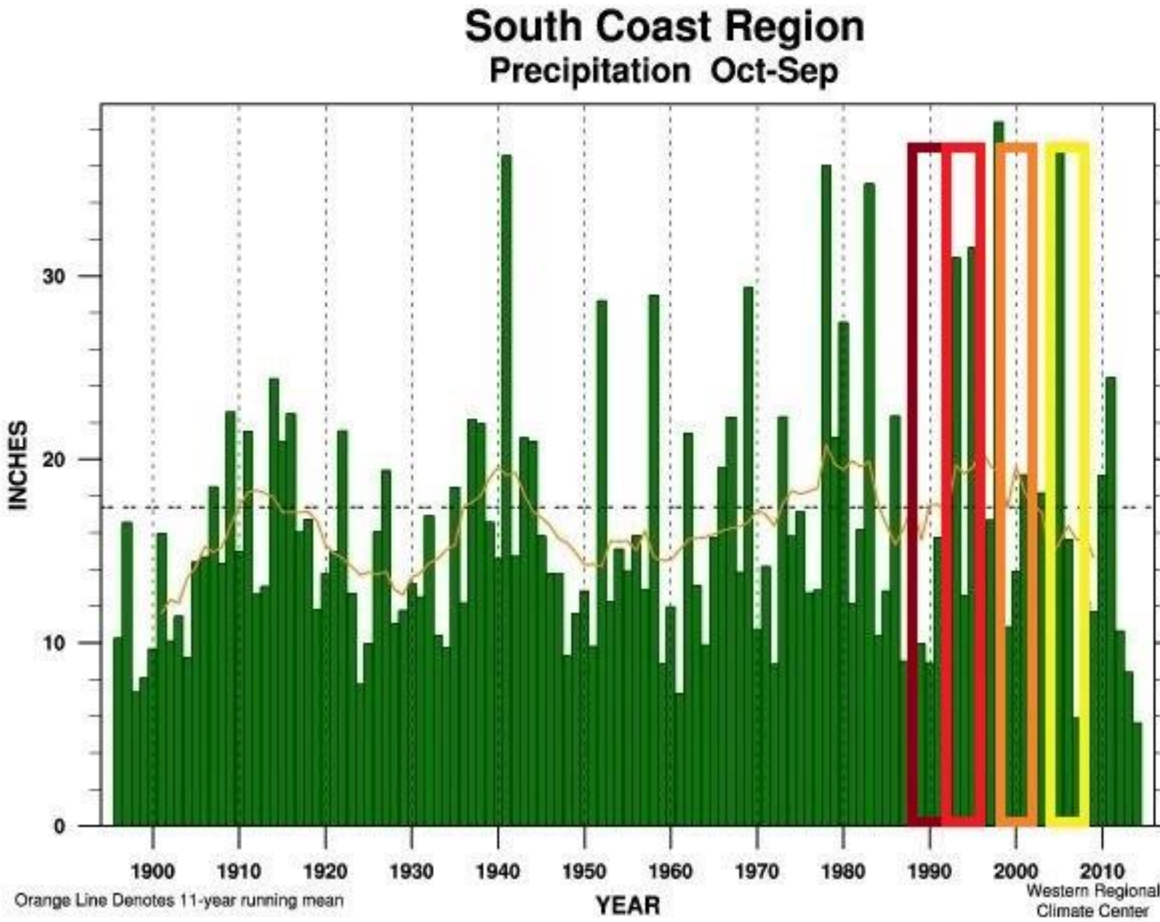

Figure C1: Regional average annual precipitation compared to the 11-year running average. The first box captures WY 1989-1991, which is characterized by below average precipitation. The second box displays WY 1993-1995 with above average precipitation. The third surrounds WY 1999-2001 and another dryer than average period, but less dry than WY 1989-1991. The final box captures WY 2005-2007 with its representative dry, average, and wet precipitation.

Table C1: Land use parameters in HEC-HMS models of altered and baseline condition: impervious data from basin delineations and verified with StreamStats (USGS, 2019c), Time of Concentration computed from basin delineations and the Kirpich Method (Kirpich, 1940), and Clark Unit Hydrograph Storage Coefficient estimated with basin delineations and according to ADOT (2014).

	Land use parameters for EFCC calibrated WY 2005-2007 models			Historical land use parameters		
	Impervious %	Time of Concentration (hr)	Clark Unit Hydrograph Storage Coefficient (hr)	Impervious %	Time of Concentration (hr)	Clark Unit Hydrograph Storage Coefficient (hr)

Chino Canyon	0.2	0.4	0.2	0	0.5	0.2
Los Coches	9.7	1.4	0.6	0	1.6	0.8
Mission	4.3	1.0	0.5	0	1.2	0.7
Mission Rocky	0.7	0.7	0.3	0	0.8	0.4
Santa Cruz	0.0	3.3	1.4	0	3.8	1.9

*Table C2: Gaged and ungaged flashiness and drying flow alteration metrics. "Gaged Modeled historical land use" is considered the most accurate, or "reference" measure of flow alteration for comparing to ungaged estimates, and is indicated by *.*

Flashiness Alteration Ratio

Site	Environmental Flow Metric	Gaged		Ungaged		
		*Modeled historical land use	Oldest three WY	Nearest neighbor regionalization (EFCC)	Nearest neighbor regionalization (NSE)	SR-HMC
Chino Canyon	RBI	1.0573	1.8397	1.5281	2.8708	1.6134
Los Coches	RBI	3.2817	1.4648	1.5364	24.8440	1.6536
Mission	RBI	1.2959	0.6214	1.2369	1.1935	1.6279
Mission Rocky	RBI	1.0511	0.9448	0.8339	0.8057	1.1092
Santa Cruz	RBI	1.0026	1.1065	0.4527	0.4488	0.6196

Flashiness Alteration

Site	Environmental Flow Metric	Gaged		Ungaged		
		*Modeled historical land use	Oldest three WY	Nearest neighbor regionalization (EFCC)	Nearest neighbor regionalization (NSE)	SR-HMC
Chino Canyon	RBI	0.0037	0.0313	0.0237	0.0447	0.0261
Los Coches	RBI	3.2817	1.4648	1.5364	24.8440	1.6536
Mission	RBI	0.0387	-0.1033	0.0325	0.0275	0.0654
Mission Rocky	RBI	0.0056	-0.0067	-0.0228	-0.0276	0.0113
Santa Cruz	RBI	0.0002	0.0061	-0.0769	-0.0781	-0.0390

Flashiness Flow Alteration (Baseline normalized)

Site	Gaged	Ungaged

	Environmental Flow Metric	*Modeled historical land use	Oldest three WY	Nearest neighbor regionalization (EFCC)	Nearest neighbor regionalization (NSE)	SR-HMC
Chino Canyon	RBI	0.0573	0.8397	0.5281	1.8708	0.6134
Los Coches	RBI	2.2817	0.4648	0.5364	23.8440	0.6536
Mission	RBI	0.2959	-0.3786	0.2369	0.1935	0.6279
Mission Rocky	RBI	0.0511	-0.0552	-0.1661	-0.1943	0.1092
Santa Cruz	RBI	0.0026	0.1065	-0.5473	-0.5512	-0.3804

Flashiness Flow Alteration (Altered normalized)

		Gaged		Ungaged		
Site	Environmental Flow Metric	*Modeled historical land use	Oldest three WY	Nearest neighbor regionalization (EFCC)	Nearest neighbor regionalization (NSE)	SR-HMC
Chino Canyon	RBI	0.0542	0.4564	0.3456	0.6517	0.3802
Los Coches	RBI	0.6953	0.3173	0.3491	0.9597	0.3953
Mission	RBI	0.2283	-0.6092	0.1915	0.1621	0.3857
Mission Rocky	RBI	0.0486	-0.0584	-0.1992	-0.2411	0.0985
Santa Cruz	RBI	0.0026	0.0963	-1.2090	-1.2280	-0.6140

Flashiness Flow Alteration (Average normalized)

		Gaged		Ungaged		
Site	Environmental Flow Metric	*Modeled historical land use	Oldest three WY	Nearest neighbor regionalization (EFCC)	Nearest neighbor regionalization (NSE)	SR-HMC
Chino Canyon	RBI	0.0557	0.5914	0.4178	0.9666	0.4694
Los Coches	RBI	1.0658	0.3771	0.4230	1.8452	0.4926
Mission	RBI	0.2578	-0.4669	0.2118	0.1764	0.4779
Mission Rocky	RBI	0.0498	-0.0567	-0.1812	-0.2152	0.1036
Santa Cruz	RBI	0.0026	0.1011	-0.7535	-0.7608	-0.4698

Drying Alteration Ratio

		Gaged		Ungaged		
Site	Environmental Flow Metric	*Modeled historical land use	Oldest three WY	Nearest neighbor regionalization (EFCC)	Nearest neighbor regionalization (NSE)	SR-HMC

Chino Canyon	< 1 cfs	0.9978	0.7816	1.2298	1.2427	1.2058
Los Coches	< 1 cfs	0.9826	0.8816	0.8576	0.8903	1.1523
Mission	< 1 cfs	0.9709	0.8007	1.0167	0.8011	1.1196
Mission Rocky	< 1 cfs	0.9870	0.9561	0.9794	0.7775	1.0642
Santa Cruz	< 1 cfs	1.0002	1.0203	0.4502	0.3333	0.7369

Drying Alteration

Site	Environmental Flow Metric	Gaged		Ungaged		
		*Modeled historical land use	Oldest three WY	Nearest neighbor regionalization (EFCC)	Nearest neighbor regionalization (NSE)	SR-HMC
Chino Canyon	< 1 cfs	-0.1624	-20.8611	13.9505	14.5783	12.7405
Los Coches	< 1 cfs	-1.3990	-10.6225	-13.1337	-9.7510	10.4537
Mission	< 1 cfs	-2.3239	-19.3356	1.2794	-19.2829	8.2959
Mission Rocky	< 1 cfs	-0.9920	-3.4663	-1.5894	-21.5810	4.5520
Santa Cruz	< 1 cfs	0.0062	0.6410	-39.3454	-64.4244	-11.5078

Drying Flow Alteration (Baseline normalized)

Site	Environmental Flow Metric	Gaged		Ungaged		
		*Modeled historical land use	Oldest three WY	Nearest neighbor regionalization (EFCC)	Nearest neighbor regionalization (NSE)	SR-HMC
Chino Canyon	< 1 cfs	-0.0022	-0.2184	0.2298	0.2427	0.2058
Los Coches	< 1 cfs	-0.0174	-0.1184	-0.1424	-0.1097	0.1523
Mission	< 1 cfs	-0.0291	-0.1993	0.0167	-0.1989	0.1196
Mission Rocky	< 1 cfs	-0.0130	-0.0439	-0.0206	-0.2225	0.0642
Santa Cruz	< 1 cfs	0.0002	0.0203	-0.5498	-0.6667	-0.2632

Drying Flow Alteration (Altered normalized)

Site	Environmental Flow Metric	Gaged		Ungaged		
		*Modeled historical land use	Oldest three WY	Nearest neighbor regionalization (EFCC)	Nearest neighbor regionalization (NSE)	SR-HMC
Chino Canyon	< 1 cfs	-0.0022	-0.2794	0.1869	0.1953	0.1707
Los Coches	< 1 cfs	-0.0177	-0.1343	-0.1660	-0.1233	0.1321

Mission	< 1 cfs	-0.0299	-0.2489	0.0165	-0.2483	0.1068
Mission Rocky	< 1 cfs	-0.0132	-0.0460	-0.0211	-0.2861	0.0604
Santa Cruz	< 1 cfs	0.0002	0.0199	-1.2215	-2.0000	-0.3573

Drying Flow Alteration (Average normalized)

		Gaged		Ungaged		
Site	Environmental Flow Metric	*Modeled historical land use	Oldest three WY	Nearest neighbor regionalization (EFCC)	Nearest neighbor regionalization (NSE)	SR-HMC
Chino Canyon	< 1 cfs	-0.0022	-0.2452	0.2061	0.2164	0.1866
Los Coches	< 1 cfs	-0.0175	-0.1258	-0.1533	-0.1161	0.1415
Mission	< 1 cfs	-0.0295	-0.2214	0.0166	-0.2208	0.1128
Mission Rocky	< 1 cfs	-0.131	-0.0449	-0.0209	-0.2503	0.0622
Santa Cruz	< 1 cfs	0.0002	0.0201	-0.7583	-1.0000	-0.3031