Development of Recommended Flow Targets to Support Biological Integrity Based on Regional Flow-ecology Relationships for Benthic Macroinvertebrates in Southern California Streams

Eric D. Stein
Raphael D. Mazor
Ashmita Sengupta
Kenny McCune
Brian Bledsoe
Stephen Adams
Sarah Eberhart
Matt Pyne
Peter Ode
Andy Rehn

Southern California Coastal Water Research Project

SCCWRP Technical Report 974

Development of Recommended Flow Targets to Support Biological Integrity Based on Regional Flow-ecology Relationships for Benthic Macroinvertebrates in Southern California Streams

Eric D. Stein[1], Raphael D. Mazor[1], Ashmita Sengupta[1], Kenny McCune[1], Brian Bledsoe[2,3], Stephen Adams[3], Sarah Eberhart[3], Matt Pyne[4], Peter Ode[5], Andy Rehn[5]

[1]Southern California Coastal Water Research Project

[2]University of Georgia, Athens

[3]Colorado State University, Fort Collins

[4]Lamar University

[5]California Department of Fish and Wildlife

SCCWRP Technical Report 974
March 2017

EXECUTIVE SUMMARY

Changes to instream flow are known to be one of the major factors that affect the health of biological communities. Regulatory, monitoring, and management programs are increasingly using biological community composition, particularly benthic invertebrates, as one measure of instream conditions, stormwater project performance, or regulatory compliance with NPDES or other requirements and regulations. Understanding the relationship between changes in flow and changes in benthic invertebrate communities is, therefore, critical to informing decisions about ecosystem vulnerability, causes of stream and watershed degradation, and priorities for future watershed management.

There are many approaches to developing flow-ecology relationships that relate hydrologic change to responses in instream biological communities that can be used to establish management targets. The Ecological Limits of Hydrologic Alteration (ELOHA) framework (Poff et al. 2010) provides a way to assess the effect of flow alteration on the condition of biological communities (vs. individual taxa) on a regional basis. Consequently, it is a useful approach for setting targets across a wide range of geographies and stream types where comprehensive detailed site-specific investigations are not practical. The framework includes elements of stream classification, estimation of flow alteration and development of flow ecology relationships based on the response of biological communities to changes in flow.

In this project, we applied to the ELOHA framework to develop regional flow-ecology relationships and targets based on responses in the benthic macroinvertebrate community. Our objectives were: 1) Develop a recommended set of flow targets for southern California streams that would maximize the likelihood of maintaining healthy biological communities as indicated by the California Stream Condition Index (CSCI) for benthic invertebrates. 2) Produce a set of

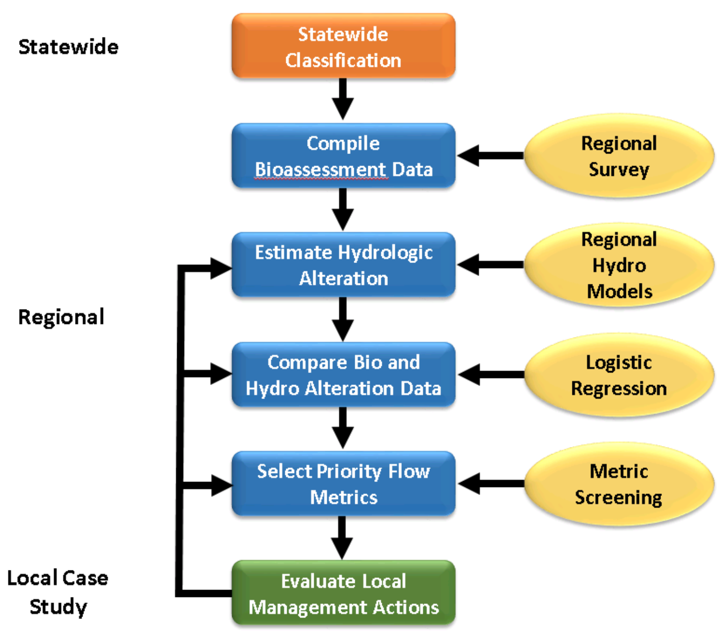

Figure ES-1. General steps used to develop flow-ecology relationships based on the ELOHA framework

tools that can be readily applied to future sites to estimate hydrologic alteration relative to biologically-define targets.

Development of the regional flow-ecology relationships relied on an ensemble of hydrologic models to estimate flow alteration at ungauged sites, and took advantage of a regional bioassessment data that allowed us to assess flow-ecology relationships at broad spatial scales. Our general approach involved developing a hydrologic classification for the entire State of California, calibrating and validating watershed models for the stream classes present in southern California, using the models to assess hydrologic change at 572 bioassessment sites, relating hydrologic change to biological responses, setting targets based on likelihood of biological response associated with changes in key flow metrics, applying the flow-ecology tools to assess regional hydrologic condition, and prioritizing sites for various management actions based on their response relative to the established flow targets (along with information on presence of other stressors; Figure ES-1).

Hydrologic Classification

Application of the ELOHA framework begins with stream classification. This is particularly important in places like California which have extreme climatic, altitudinal, and geologic gradients which affect stream morphology, flow patterns, and biological communities. This complexity, combined with spatially variable patterns of land use (e.g. urban, agricultural, timber, hydropower) produces highly variable flow responses in streams that must be accounted for during development of flow management targets. We developed a statewide classification system using a two-step approach: First, we classified stream reaches according to watershed characteristics. Streams were clustered into 7 classes based on differences in winter precipitation, geology, soil characteristics, and mean watershed elevation (Figure ES-2). Second, we used flow data from a subset of streams with gauge information to test and refine the stream classes and determine which hydrologic variables best separate streams into their respective classes. Most of the hydrologic variables that best discriminated between-stream typology classes were indicative of high flows, mean or median flows, or flow timing.

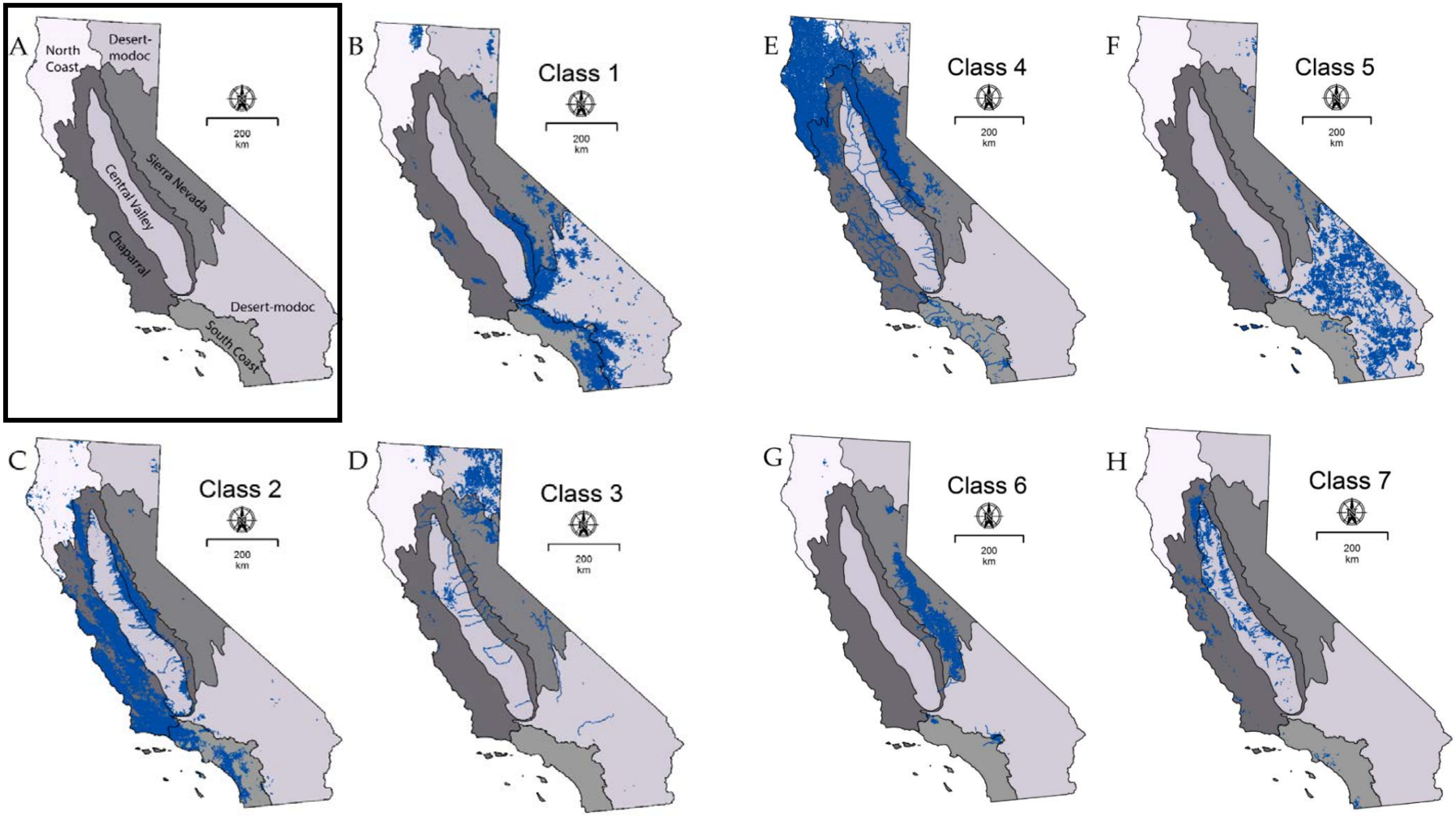

Figure ES-2. Maps showing the distribution of each stream class in the state. The Perennial Stream Assessment (PSA) regions are defined in panel A, with each of the 7 classes shown in panels B-I. Stream segments assigned to each class are represented by the blue lines.

Assessing Hydrologic Alteration

We modeled hydrologic alteration at 572 ungauged bioassessment sites in Southern California using a two-step approach. First, we developed an ensemble of 26 calibrated and validated watershed models applicable to all 572 sites. One of the 26 models was assigned to each of the 572 ungauged sites based on similarities in catchment properties. Second, we used the models to generate hourly flows for both current and reference conditions at each site and used the flow data to calculate a suite of flow metrics that represent different hydrograph components (e.g. magnitude, duration, frequency). Reference conditions were simulated by adjusting the current models to reflect pre-urbanization conditions by setting imperviousness to zero to mimic no urban land use, and by increasing initial losses to account for greater land availability.

Using a suite of over 100 flow metrics, we estimate that approximately 79% of the region shows some degree of hydrologic alteration, and approximately 40% of the sites can be considered severely altered with at least 10 metrics exhibiting severe hydrologic alteration. Among the five metric categories (timing, frequency, magnitude, duration and variability), the most common alteration was an augmentation of the magnitude metrics. In general, hydrologic alteration is pervasive in catchments with total impervious cover higher than 2% (Figure ES-3).

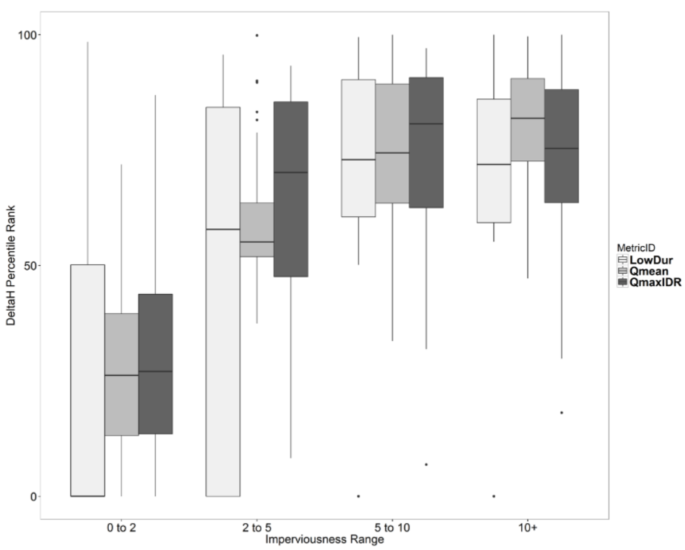

Figure ES-3. Hydrologic alteration for three selected flow metrics LowDur (duration), Qmean (magnitude), and QmaxIDR (variability) at various levels of total impervious cover.

Establishing Flow-Ecology Targets and Evaluating Regional Condition

We developed recommendations for flow-ecology targets by estimating the probability of healthy biological condition, based on the California Stream Condition Index (CSCI), as a function of different levels of hydrologic alteration. Targets were set at the level of hydrologic alteration corresponding to a 50% decrease in the probability of healthy biological conditions. We prioritized a subset of seven flow metrics based on the strength of their relationship with changes in the biological community composition, as well as the following criteria (Table ES-1):

- Ability to differentiate reference sites and non-reference sites
- Strong relationship to biological condition based on boosted regression tree analysis and can produce a hypothesized ecological response
- Ability to be modeled under both current and reference conditions with a high level of confidence
- Amenability to management actions, with predictable responses to changes in flow conditions
- Representation of different components of the hydrograph (e.g. magnitude vs. duration)
- Minimal redundancy with other metrics

The seven priority flow metrics were aggregated into an overall hydrologic alteration index and applied to the regional bioassessment data set. This index indicates where hydrology is altered to a level associated with undesirable changes in the instream biological community (as indicated by composition of benthic macroinvertebrates). Approximately two-thirds of stream-kms were considered minimally altered based on our criteria (i.e., hydrologic alteration index score of zero). Where alteration occurred, it was most extensive in urban streams (91%), followed by agricultural streams (80%). Alteration was limited to only 11% of stream-kms draining undeveloped catchments (Figure ES-4). Magnitude metrics (particularly those associated with high flows) and variability metrics showed the greatest influence on biological response variables. In contrast, timing metrics had relatively little influence on biological response.

Figure ES-4. Hydrologic alteration scores at sites in the Southern California region. Urban areas are represented as dark gray. Boundaries of major hydrologic regions are shown.

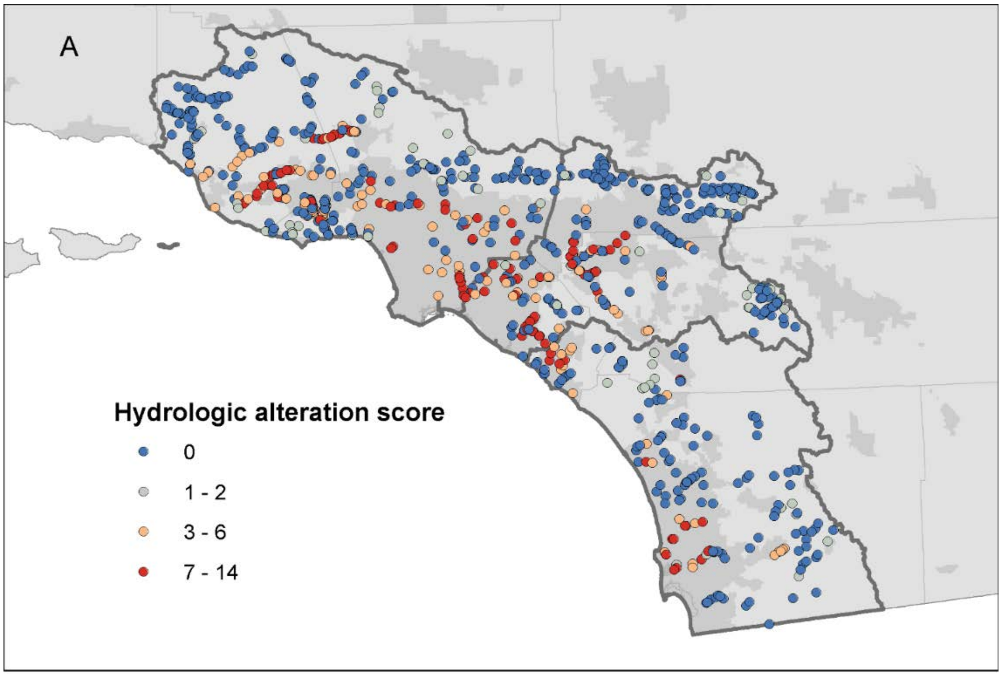

Table ES-1. Priority hydrologic metrics and associated thresholds used in the regional flow-ecology relationships. Metrics are grouped the hydrograph component they represent. Thresholds are expressed as the change in metric value associated with poor biologic condition (CSCI <0.79). Thresholds can be based on increasing or decreasing flows. Metric effects on biology were typically strongest during either average, wet, or dry rainfall years, or all years combined (overall). NT= no threshold established.

Hydrograph Component	Metric	Metric Definition	Critical precipitation condition	Decreasing Threshold	Increasing Threshold
Duration	NoDisturb (days)	median annual longest number of consecutive days that flow is between the low and high flow threshold	Average	-64	NT
	HighDur (days/event)	median annual longest number of consecutive days that flow was greater than the high flow threshold	Wet	-3	24
Magnitude	MaxMonthQ (m3/s)	Maximum mean monthly streamflow	Wet	NT	1.5
	Q99 (m3/s)	streamflow exceeded 99% of the time	Wet	-0.01	32
Variability	RBI (unitless)	Richards-Baker index of stream flashiness	Dry	NT	0.25
	QmaxIDR (m3/s)	Interdecile range of flow	Overall	-5	2.5
Frequency	HighNum (events/year)	median annual number of events that flow was greater than high flow threshold	Dry	NT	3

By applying a regionally applicable ensemble of hydrologic models to a large bioassessment data set, we were able to model biological responses across a wide range of range of conditions, and derive flow targets that can be applied to sites throughout the region. Furthermore, because these targets are based on probabilities of biologic response to varying degrees of hydrologic alteration (vs. static targets) managers can adjust the targets according to their tolerance for risk or based on the importance of competing water demands. Development of regionally applicable flow targets allows application to any bioassessment site and reduces the need to develop local flow–ecology relationships for every stream of interest, as is the case in more traditional instream flow methods. This provides a mechanism for prioritizing management actions based on consistent flow-ecology relationships. The tools developed through this project are readily transferable for local stakeholders to produce measures of hydrologic change for any location of interest, and to explore how those values would change under different land-use or management scenarios.

ACKNOWLEDGEMENTS

We thank our technical advisory and stakeholder workgroups for their participation throughout this project. Their input and critical review improved both the technical quality and the management applicability of this work. Our partners from the U.S. Geological Survey, Daren Carlisle, Chris Konrad, Larry Brown, and Jason May, provided key technical input and assistance on modeling to support classification and development of the flow-ecology relationships. Funding for this project has been provided in full or in part through an agreement with the State Water Resources Control Board under Grant Agreement #12-430-550. The contents of this document do not necessarily reflect the views and policies of the State Water Resources Control Board, nor does mention of trade names or commercial products constitute endorsement or recommendation for use."

SUGGESTED CITATION

Stein, E.D., R.D. Mazor, A. Sengupta, K. McCune, B. Bledsoe, S. Adams, S. Eberhart, M. Pyne, P. Ode and A. Rehn. 2017. Development of Recommended Flow Targets to Support Biological Integrity Based on Regional Flow-ecology Relationships for Benthic Macroinvertebrates in Southern California Streams. Technical Report 974. Southern California Coastal Water Research Project. Costa Mesa, CA.

ELECTRONIC SUPPLEMENTAL MATERIALS

The following supplemental materials are available electronically upon request. Please contact pubrequest@sccwrp.org to request a copy.

1. GIS files for stream classification
2. Ensemble model: contains six sub-folders and a Word document with detailed instructions. All of the precipitation files used for model inputs also are available.
3. Output of models: daily flows for current and reference conditions for 572 bioassesment sites
4. Logistic regression plots and thresholds
5. Flow metric values for 572 bioassessment sites, biological endpoints and whether the sites pass or fail designated thresholds (three files)

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

1.0 INTRODUCTION

Flow regime has been shown to affect a broad suite of ecological processes and biological communities (Bunn and Arthington 2002, Naiman et al. 2002, Poff et al. 1997, Poff and Zimmerman 2010, Novak et al. 2015). Many studies have demonstrated that alterations of flow regime can be associated with changes in macroinvertebrate assemblages, which are used as key bioindicators for many regulatory and management programs globally (Pringle et al. 2000, Miller et al. 2007, DeGasperi et al. 2009, Poff & Zimmerman 2010). Although a basic understanding of the relationship between flow alteration and ecological response exists (Poff et al. 2010), few studies have demonstrated how to develop regulatory or management objectives (or targets) based on these relationships. Establishing quantitative and predictive relationships between change in flow and change in biological community composition is a critical step in understanding factors that contribute to reduced biological and condition and to using bioassessment indicators to establish measures of project performance or regulatory compliance.

Various approaches have been used to develop relationships between flow characteristics and biological response. Examples include use of habitat suitability models that relate flow change to requisite habitats for target taxa (e.g., MesoHABSIM, Parasiewicz 2009; and PHABSIM, Beecher et al. 2010); establishment of functional flow regimes to support species of management concern (McClain et al. 2014, Yarnell et al. 2015); and use of statistical ranges of sustainability based on unaltered hydrographs (Richter et al. 2011). Concepts from several of these approaches have been organized into the Ecological Limits of Hydrologic Alteration (ELOHA) framework (Poff et al. 2010). The ELOHA framework uses a variety of hydrologic and biologic tools to determine and implement environmental flows at the regional scale. Results of the ELOHA analysis can inform management decisions, such as release rates from dams, reservoirs or basins, diversion volumes for irrigation or water re-use, or flows associated with stream restoration. Because the ELOHA framework provides a way to assess the effect of flow alteration on the condition of biological communities (vs. individual taxa) on a regional basis, it is a useful approach for setting targets across a wide range of geographies and stream types where comprehensive detailed site-specific investigations are not practical.

The ELOHA framework establishes targets by comparing changes in hydrology (based on specific flow metrics) to changes in biology (based on the composition of target biological communities). The framework includes elements of stream classification, estimation of flow alteration (termed “delta H”) and development of flow ecology relationships based on the relationship between delta H and changes in the biological community (“delta B”). Implementing the ELOHA framework involves addressing the following technical challenges:

1. California is a complex state with continental-scale variability in natural conditions. Streams are expected to respond differently to flow alteration as a function of their setting and underlying hydrological processes. Grouping sites into similar categories based on their native properties can reduce “noise” in the flow-ecology relationships associated with this natural variability
2. Development of regional targets requires having many sites with data of both biological and hydrological alteration. Fortunately, California’s robust regional monitoring programs and use of predictive indices of condition provide thousands of sites with data

on biological alteration. Estimating hydrologic alteration for each of these sites requires development of tools or models that can readily be applied to large numbers of sites in an efficient manner.

3. There are hundreds of potential flow metrics that can be calculated and numerous ways that thresholds of response can be determined. Ultimately, a subset of metrics and thresholds should be selected based on considerations of ecological relevance, relevance to management actions, and parsimony. A robust metric screening process can help develop a recommended set of metrics and thresholds that can be used to produce effective and efficient flow targets.

In this project, we developed approaches to address each of the above challenges for stream in Southern California. Our objectives were: 1) Develop a recommended set of flow targets for southern California streams that would maximize the likelihood of maintaining healthy biological communities as indicated by the California Stream Condition Index (CSCI) for benthic invertebrates. 2) Produce a set of tools that can be readily applied to future sites to estimate hydrologic alteration relative to biologically-define targets. 3) Demonstrate application of the tools in a specific location to assess condition and evaluate management scenarios. Our general approach involved a sequence of steps designed to meet the study objectives at different scales ranging from statewide for the hydrologic classification, to regional for assessing hydrologic alteration and establishing putative flow recommendations necessary to protect stream health, to watershed-scale for the demonstration case study (Figure 1-1).

In this report we summarize three major products that were developed to address the technical challenges listed above and our overall project objectives: 1) a classification system aimed at reducing "noise" in flow-ecology relationships associated with natural variability between stream sites, 2) a hydrologic modeling framework and tools that can be used to assess degree of hydrologic alteration at any site in southern California and predict likely future hydrologic changes associated with management actions, and 3) a set of priority flow metrics and thresholds that are associated with reduced biological condition as indicated by CSCI. Results of the demonstration case study are provided in a companion technical report (SCCWRP TR # 948).

In addition to this summary, the following data products have been produced as part of this project:

1. Statewide hydrologic classification system – GIS layer attributed to NHD
2. Ensemble of 26 watershed models
3. Estimates of hydrologic alteration for ca. 800 bioassessment sites in S. Ca. based on ca. 40 flow metrics – this includes estimates of current and historical/natural flow values for all metrics
4. Logistic regression outputs that related CSCI and component metrics to each flow metric.
5. Ratings for each S. Ca. bioassessment site in relative to established flow targets

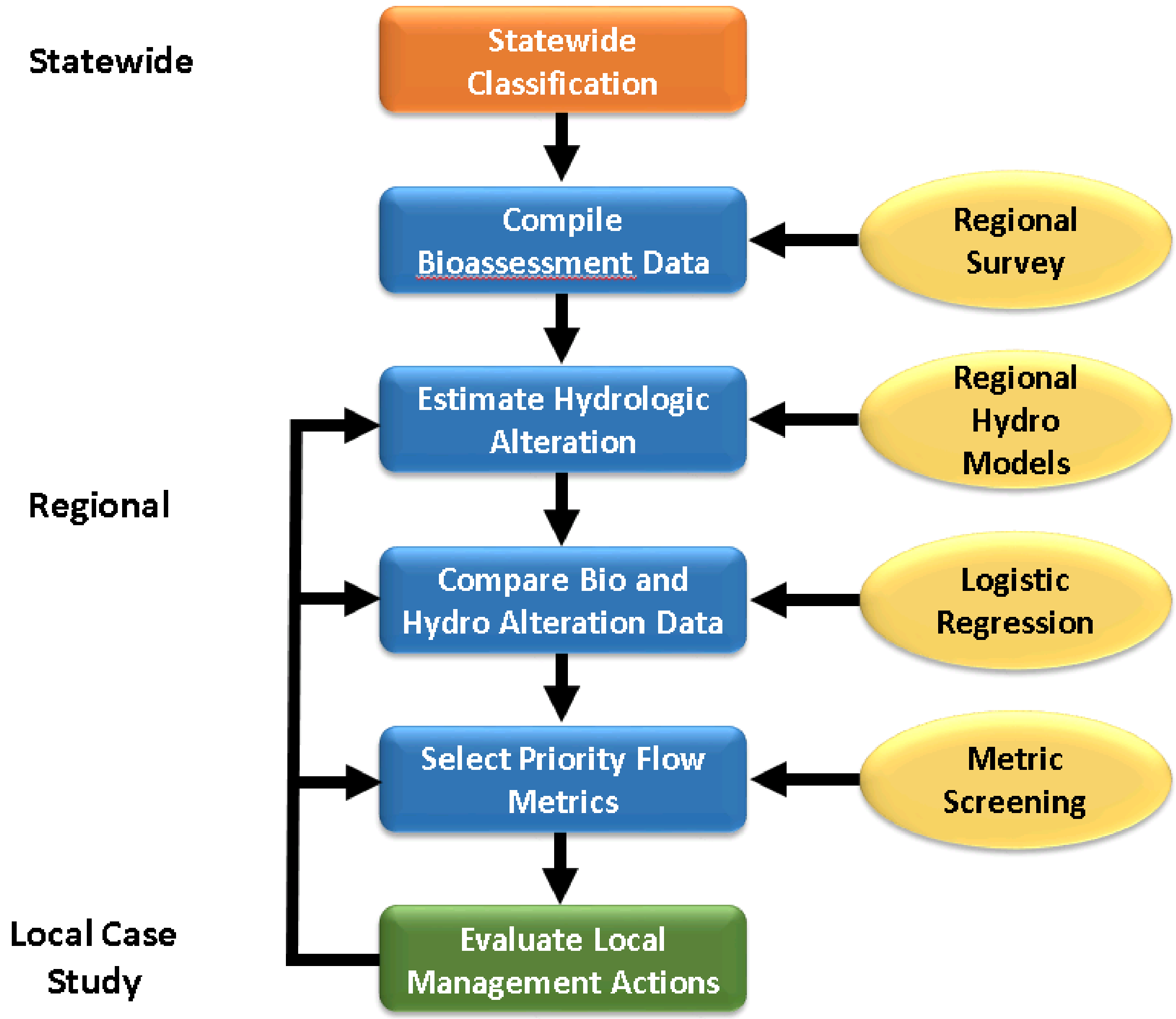

Figure 1-1. General steps used to develop flow-ecology relationships based on the ELOHA framework

2.0 Statewide Hydrologic Classification

2.1 Background

Classifying streams into relatively homogenous groups based on their hydrology and geomorphology is a foundational step in the ELOHA framework. Geomorphic and hydrological classification is important for ecological flow assessments because it provides a spatially explicit understanding of flow regime variation among rivers and regions and how a stream's natural characteristics and setting influence the relationship between flow and biology (Kennard *et al.*, 2010, Poff *et al.*, 2010). Different stream types may be expected to respond differently to flow alteration (i.e. may be more or less sensitive to different alterations). Robust classification will improve the development of flow-ecology relationships by reducing some of the noise in relational models associated with normal landscape variability (Brown et al. 2014).

Hundreds of flow classification studies have been completed over the past 30 years using a variety of approaches (Belmar et al. 2011). Olden et al. (2012) reviewed more than two dozen recent approaches to hydrologic classification and divided the approaches broadly into 1) deductive approaches that classify streams based on physical and environmental factors such as catchment properties and rainfall patterns, and 2) inductive approaches that classify flow regimes based on analysis of stream gauge networks and other flow data. Olden et al. (2012) recommend that deductive approaches should be used when the goal is to provide a general description of hydrologic patterns based on first principles and/or when streamflow data are limited. This approach is also geographically independent, classifying streams based on similar environmental conditions regardless of location.

We employed a hybrid deductive-inductive approach to classify all stream reaches in the State of California. Our goal is to produce a comprehensive hydrologic classification that groups streams into similar inherent flow properties across California's climatically and geologically diverse landscape. We start with a deductive classification, followed by an inductive approach using the statewide stream gauge network to validate and refine class membership and to identify hydrologic variables that best differentiate basins with minimal anthropogenic disturbance (hereafter designated as "reference" streams or sites) from streams altered by anthropogenic disturbances such as alterations to the natural hydrological regime, land use changes, or pollution (hereafter designated as "non-reference"). This approach will support regional flow-ecology analysis by helping to identify a priority set of flow variables that should be prioritized for subsequent analysis of relationships to biological endpoints, per the ELOHA approach.

2.2 Methods

2.2.1 Classification analysis (deductive analysis)

All stream reaches (i.e., segments from the National Hydrography Dataset (NHD)) for California served as the base layer for classification modeling (ca. 135,000 segments; 1:100,000-scale NHDPlus v.1; Figure 2-1). Statistical models of monthly flows are being developed for California streams for an unrelated project (Carlisle et al. In Review), and provided insight into the selection of physical attributes that may affect streamflow patterns. Each stream segment was attributed with approximately 150 physical characteristics assigned at the drainage area scale. These characteristics included topography, geology, soil type, and long-term (1950-2000) average precipitation and other climatic attributes. The physical characteristics most influential to streamflows (Table 2-1) were selected from the larger set based on examination of relevant literature and the application of first principles of watershed hydrology (i.e. flow is affected by watershed size, slope, maximum elevation, annual precipitation, etc.) during the hydrologic modeling work in Carlisle et al. (in review). Classification of segments was then performed on the selected subset of physical characteristics. Unsupervised Bayesian mixture modeling (BMM) was used to perform the classification of each stream segment based on physical attributes of the watershed. BMM has several advantages for hydrological classification (Webb et al., 2007; Kennard et al. 2010), with the number of classes being determined objectively and each individual segment may have membership in each class with known probability. This ability to classify observations into multiple groups better reflects the natural world, because there are many streams that may be transitional from one distinct type (e.g., snow melt perennial) to another (e.g., snow melt-winter rain perennial). The procedure iteratively modeled the probability that the observations belong to pre-defined probability distributions of the explanatory variables (i.e., physical watershed attributes) and then selected the most parsimonious solution. Calculations were carried out in AutoClass@IJM (Achcar et al., 2009). One distinct disadvantage using an objective classification technique is that it often produces a number of classes so large that they would be difficult to implement into ecosystem management applications. To further reduce the number of classes produced by the BMM, we further grouped our classes into *meta-classes* following an approach similar to that taken by Brown et al. (2014). The aggregation involved calculating the mean value of each watershed attribute variable across sites within each of the BMM classes. For each stream class, we correlated the mean and median values of each environmental variable. The correlations were all >0.9, which suggests that the distribution of data were not highly skewed and that the mean is an adequate measure of central tendency in each class. A hierarchical classification (Ward's linkage on Euclidean distances) was conducted on the BMM classes based on mean values for watershed attributes. To determine the number of final classes in the hierarchical clustering, we calculated Van Sickle and Hughes (2000) classification strength for 2-13 classes and subjectively evaluated the classes across the Californian landscape from a potential managerial perspective. The use of an additional classification technique does add additional noise, but allowed us to reduce our number of classes to a number that could be applied to stream management decisions.

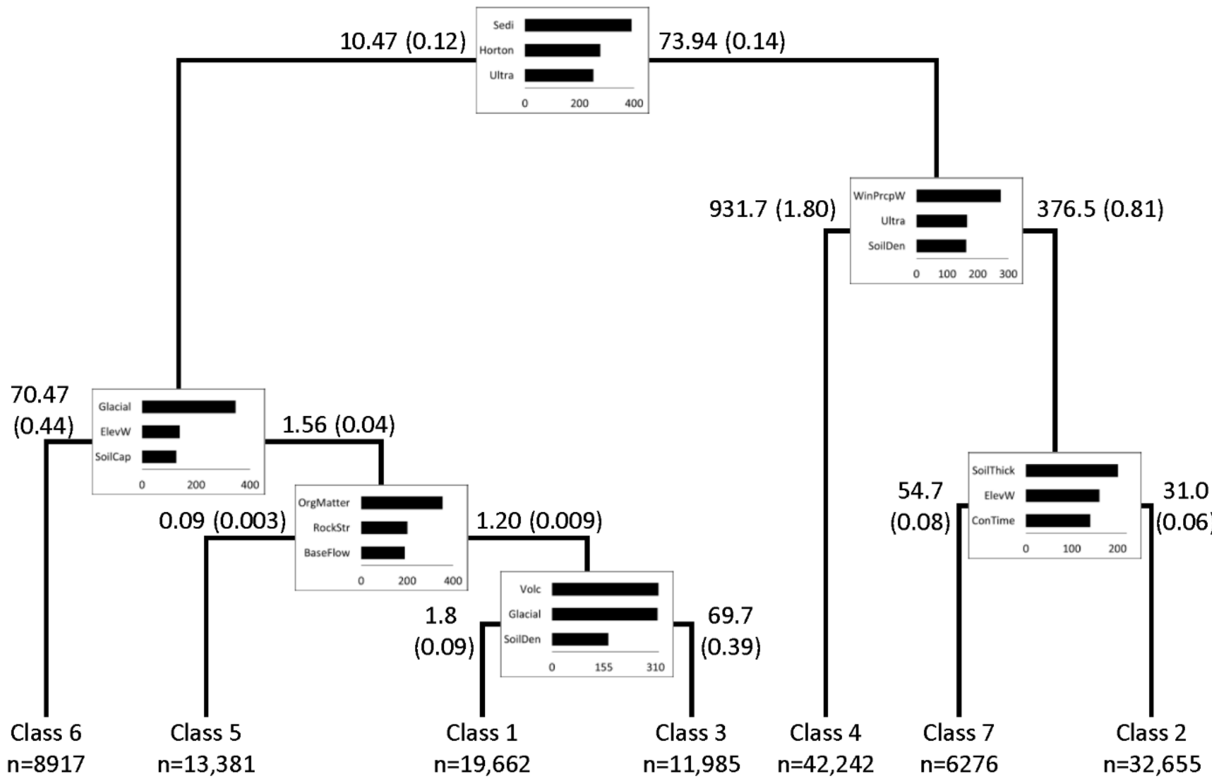

Figure 2-1. Flow diagram showing the type of datasets, explanatory variables, important questions, and statistical analyses used in this paper. The datasets and variables are represented by rectangles with rounded corners while the analyses are represented by the rectangles with square corners.

Table 2-1. Climatic and physical attributes of watersheds used to classify stream segments in California (after Falcone, 2012).

Variable	Description	Data source
ElevW	mean elevation of watershed (m)	National Elevation Dataset (NED)
ElevC	mean elevation of reach catchment (m)	NED
Elevdiff	difference in mean elevation between catchment and watershed (m)	NED
SlopeW	mean slope of watershed (%)	NED
SlopeC	mean slope of reach catchment (%)	NED
DrainA	delimited area of watershed (km^2)	NED
WinPrcpW	mean precipitation of the watershed in October-March (cm)	Parameter elevation Regression on Independent Slopes Model (PRISM)
WinPrcpC	mean winter precipitation of the reach catchment	PRISM
WinPrcpR	Ratio of mean winter precipitation of the reach catchment to watershed	PRISM
SumPrcp	percent of annual precipitation falling in June-September	PRISM
Evapo	mean evapotranspiration within watershed (mm/year)	PRISM
Dunne	percent of streamflow composed of overland flow	NED, U.S. General Soil Map (STATSGO)
Horton	percent of streamflow composed of Horton flow	NED, STATSGO
TopoWet	index of topographic wetness (ln(m))	NED, STATSGO
ConTime	subsurface flow contact time (days)	NED, STATSGO
BaseFlow	percent of streamflow composed of groundwater	
SoilB	soil hydrologic group B (%)	STATSGO
SoilC	soil hydrological group C (%)	STATSGO
SoilCap	soil water capacity (unitless)	STATSGO
SoilPerm	soil permeability (in/hour)	STATSGO
SoilDen	soil bulk density (g/cm3)	STATSGO
SoilThick	Mean total soil thickness examined (in^2)	STATSGO

OrgMatter	Soil with organic matter content (%)	STATSGO
RockStr	rock compressive strength (MPa)	(Olson and Hawkins, 2012)
Sedi	percent of watershed composed of sedimentary bedrock	(Reed and Bush, 2005)
Ultra	percent of watershed composed of ultramafic bedrock	(Reed and Bush, 2005)
Volc	percent of watershed composed of volcanic bedrock	(Reed and Bush, 2005)
Glacial	percent of surficial rock composed of mountain glacial deposits	(Reed and Bush, 2005)

We tested the validity of our stream classes using two statistical techniques. First we used permutation-based analysis of variance (PERMANOVA) to determine if our stream typology classes were significantly different from each other using the physical and climate characteristics of watersheds. Second, we used random forests to determine which physical and climatic characteristics best explain the difference between groups for each split in the hierarchical clustering process. PERMANOVA is a non-parametric, multivariate analysis of variance that tests for differences between groups using a similarity measure (e.g., Euclidean distance) between groups and constructs a p-value by randomly permuting group values (Anderson 2001; Anderson and Walsh 2013). The PERMANOVA analysis is robust, but is sensitive to heterogeneity of variances between groups if groups varied in size. We developed a balanced design approach where an equal number of sites were selected from each group. PERMANOVA is also computationally intensive, and we could not perform the analysis with all ~135,000 sites. Instead, we randomly selected 1000 sites from each class for this analysis. We performed a global and then pairwise PERMANOVAs for all classes using the 'vegan' package (Oksanen et al. 2016) in R (R Core Team 2016) with 999 permutations.

Random forest is a powerful tool in determining which variables best account for differences between groups (McCune and Grace 2002). Random forests are an ensemble of classification and regression tree analyses (CART) using a bootstrap technique, where 70% of the observations are randomly assigned to each CART tree and results from each tree are averaged (Breiman, 2001). CART is a nonparametric technique that uses explanatory variable values to dichotomously partition sites into increasing homogenous groups, forming a tree-like sequence of partitions (McCune and Grace 2002). The accuracy of the random forest analysis was measured by the out-of-bag error rate (OOB). The OOB is calculated for each tree by measuring the misclassification of the 30% of sites not included in the bootstrap sample (Breiman 2001). Additionally, the importance of each explanatory variable can be estimated using the mean decrease of accuracy, where each variable is permuted and OOB re-estimated (Cafri 2013). Explanatory variables with high positive importance notably reduce OOB, while some variables can have a negative mean decrease of accuracy, meaning that the inclusion of the variable produces higher error rates than by chance alone. We sequentially removed all explanatory variables from the analysis that produced negative importance values by removing variable with the largest negative value and rerunning the analysis. We did this until no negative importance value remained.

Random forest (RF) is a robust, non-parametric classification technique, but some biases can be incorporated into the model design. First, if a dataset is unbalanced, with one class having a substantially greater number of sites, then RF will bias correct classifications towards the class with greater numbers (non-reference sites in our case). The effectiveness of RF is assessed using misclassification rates and if a dataset is unbalanced, then RF can classify most sites into the largest class and still keep the misclassification rate low. To counter this bias, we selected a random sample of non-reference sites equal to the number of reference sites for each tree. Another issue with RF is that an explanatory variable's importance is inflated if it is correlated with other explanatory variables (Strobl et al. 2009, Cafri 2013). One solution was to use a conditional permutation importance measure (Strobl et al. 2009), but this method consumes massive amounts of computing memory and can only be implemented with a few explanatory variables. An alternative method was to increase the number of explanatory variables sampled for each split in each tree. This increased the chance that the best variable will be selected versus a correlated, yet inferior variable. This method also increased the misclassification rate (Cafri 2013), so we performed our analyses with a low sampling number (5) and a high sampling number (all explanatory variables) and found that a higher sampling number increased the misclassification error by only 1-2%, which we found to be acceptable. We should also note that this increase in number of variables sampled per split means that the variation between trees in the RF is mostly due to bootstrap sampling of sites for each tree (Cafri 2013). We performed this balanced design, large split-sample RF analysis using the 'randomForest' package (Liaw and Wiener 2002) in R with 5001 trees.

2.2.2 Determination of reference vs non-reference status

The hydrologic validation of our classes (step 2) required reaches with hydrologic data and a known reference/non-reference status for the watershed of each reach. We computed hydrologic metrics for reaches with USGS gauges, but the reference/non-reference status of many gauges are unknown, so we developed a random forest model to predict the reference/non-reference status of the unknown gauges. We began with a data set consisting of 1305 U.S. Geological Survey stream gauges across California. Within this data set 600 gauges had been previously identified as non-reference and 154 identified as reference using 4 criteria: 1) geospatial measures of watershed-scale disturbance (Figure 2-2), 2) visual screening of topographic maps for stream alteration, 3) local expert judgment, and 4) detection of water extraction using ≥20 years of daily discharge data (Falcone et al. 2010). The remaining 551 gauges had not been previously classified by Falcone et al. due to a lack of sufficient discharge data or lack of local reports on discharge regulation and diversion. For these gauges, we classified them as reference or non-reference by calibrating a random forest classification model using the sites with known reference/non-reference status and estimating the status for the unknown sites. Specifically, we 1) derived various disturbance variables (Table 2-2), including most disturbance metrics used by Falcone et al. (2010), for all gauges via a geographic information system (GIS; ArcMAP™ 10.2.2, ESRI, Redlands, CA), 2) calibrated a random forest classification tree analysis using the 600 known reference/non-reference sites, with reference status as the classification and the disturbance variables as the explanatory variables and 3) predicted the reference/non-reference status of the unknown gauges by entering their disturbance variable values into the calibrated random forest model. We performed this analysis using the balanced design, large split-sample RF analysis described above.

Table 2-2. Watershed-scale disturbance variables used to distinguish reference and non-reference gage sites.

Variable	Description	Unit	Data source
WArea[a]	Watershed area	km^2	NHD plus (v. 1), GAGESII, or derived from 1-arcsecond DEMs
PerrStream	Perennial streams	km/km^2	NHD plus (v. 1)
Imp	Mean impervious land area	%	National land cover database (NLCD), 2011
PopDens	Population density	$\#/km^2$	US Census Bureau, 2010
RoadDens	Total road Length	%	US Census Bureau, 2009
IrrAg	Irrigated agriculture land use	%	USGS, 2012
Crops	Crops land use	%	NLCD, 2011
Pasture	Pasture land use	%	NLCD, 2011
Devl	Developed land use	%	NLCD, 2011
HDevl	High intensity developed land use	%	NLCD, 2011
Shrub	Shrub land use	%	NLCD, 2011
Grass	Grassland land use	%	NLCD, 2011
Forest	Forest land use	%	NLCD, 2011
BLMGraz	BLM grazing allotments	%	BLM
ProtectLand	Federal, state, and non-profit lands with land use restrictions	%	States of California, Oregon, Nevada; NPS, BLM, FWS, FS
Mine	Mine and mineral plants	$\#/km^2$	USGS: mineral resources, 2003

NPDES	Pollutant discharge (NPDES) sites	#/km^2	EPA: geospatial data access, 2009
Dam	Number of dams	#/km^2	USACE: national inv. of dams, 2009
DamVol	Total dam volume	Acre-feet/km^2	USACE: national inv. of dams, 2009
Canal	Canals	km/km^2	NHD plus (v. 1)
Pipes	Pipelines	km/km^2	NHD plus (v. 1)
StrmRdInter	Road-stream intersections	#/km^2	NHD plus (v. 1), US Census Bureau
FrshWith	Freshwater withdrawals	Mgal/day/km^2	USGS: water use in US, 2010

[a]Watershed area (WArea) is not a direct measure of disturbance, but can have an effect on the spatial arrangement and influence of disturbance (King *et al.*, 2005)

2.2.3 Hydrologic validation of class membership (inductive analysis)

A series of flow variables with presumed ecological relevance were calculated for all gaged stream segments in California (Table 2-3). Konrad et al. (2008) found that variables of streamflow variation at daily to inter-annual scales were among the most common characteristics associated with limits on invertebrate assemblages. Flow variables were selected based on four criteria: 1) capacity to capture the range of natural hydrologic variability, 2) known ecological relevance and sensitivity to regional styles of flow alteration, 3) prediction accuracy (confidence) of models for a given variable, and 4) interpretability and utility in a management context. We included raw variables and also included streamflow metrics (m^3/s) that had been standardized according to watershed area. Flow variables were calculated for gauges with at least 5 years of flow data in the last 15 years (2000-2014). The use of only 5 years of flow data represents a distinct tradeoff between the ability to include more sites, particularly reference-condition sites, versus increased power to detect variation in hydrology metrics between classes with greater year of flow. We felt that the need for more sites, particularly in classes with few reference condition reaches, superseded the increased power.

Table 2-3. Flow Variables used to validate stream class membership and separate reference from non-reference streams in each class.

Variable	Description
Qmean	mean streamflow for the period of analysis (m^3/sec)
Qmedian	median annual mean streamflow (m^3/sec)
QmeanIDR	90th percentile of annual mean streamflow - 10th percentile of annual mean streamflow)/50th percentile of median annual mean streamflow
Qmed	median daily streamflow (m^3/sec)
Qmax	median annual maximum daily streamflow (m^3/sec)
QmaxIDR	90th percentile of annual maximum streamflow - 10th percentile of annual maximum streamflow)/50th percentile of annual maximum streamflow
HighNum	median annual number of events that flow was greater than high flow threshold, an event is a continuous period when daily flow exceeds the threshold (events/year)
HighDur	median annual longest number of consecutive days that flow was greater than the high flow threshold (days/event)
Qmin	median annual minimum daily streamflow (m^3/sec)
QminIDR	90th percentile of annual maximum streamflow - 10th percentile of annual maximum streamflow)/50th percentile of annual maximum streamflow
LowNum	median annual number of events that flow was less than or equal to the low flow threshold, an event is a continuous period when daily flow was less than or equal to the threshold (events/year)
LowDur	median annual longest number of consecutive days that flow was less than or equal to the low flow threshold (days/event)
NoDisturb	median annual longest number of consecutive days that flow between the low and high flow threshold (days)
Hydroperiod	fraction of period of analysis with flows
FrYrNoFlow	fraction of years with at least one no-flow day
MdNoFlwDay	median annual number of no-flow days (days/year)
Less1CFS	Fraction of time period with flows less than 1 ft^3/sec (0.0283 m^3/sec)
RBI	Richards-Baker flashiness index, a measure of abrupt changes in flow over short periods of time (i.e., flashiness). The absolute values of daily flow differences divided the yearly sum of daily flows (%)
PDC50	median percent daily change in streamflow, no flow days are not included (0.01 = 1%); (%)

Variable	Description
SFR	the 90th percentile of percent daily change in streamflow on days when streamflow is receding (a measure of storm-flow recession; %)
BFR	the 50th percentile of percent daily change in streamflow on days when streamflow is receding (a measure of base-flow recession; %)
MaxMonth	month of maximum mean monthly streamflow
MxMnthQ	maximum mean monthly streamflow (m^3/sec)
MinMonth	month of minimum mean monthly streamflow
MnMnthQ	minimum mean monthly streamflow (m^3/sec)
Q01, Q05, ...,Q95,Q99	streamflow exceeded 1%, 5%, 10%, 25%, 50%, 75%, 90%, 95%, and 99% of the time (m^3/sec)
S_Qmean, S_Qmedian,…,S_Q99	The following flow variables were also standardized by their watershed area and included in the dataset: Qmean, Qmedian, Qmed, Qmax, Qmin, MxMnthQ, MnMnthQ, and Q01-Q99 (m^3/sec/km^2)

The hydrology metrics were used to validate and refine the stream typology classes in two ways. First, we compared the ability of the hydrology metrics to differentiate the watershed typology classes. We were limited in this analysis to the 139 reference-condition, gaged reaches. We used balanced-design pairwise PERMANOVAs, with the number of sites selected for both classes equal to the number of reference sites in the smaller class, to determine if hydrology metrics can significantly differentiate classes. We followed with pairwise RF analyses to determine which hydrology variables best differentiate classes. Second, we used the PERMANOVA and RF analyses to determine if hydrology variables can significantly separate reference condition streams from non-reference condition streams in each of the stream classes and determine which hydrology metric best differentiate reference from non-reference streams in each class. For this analysis we included the 138 reference sites plus 479 non-reference sites. We could not include gauges from class 7 due to only one reference condition site in that class.

2.3 Results

2.3.1 Identification of stream typology for California

The unsupervised classification model produced 109 statistically distinct classes among the ~135,000 stream segments. Because this was too many to interpret, we used the secondary hierarchical classification to aggregate these 109 classes into seven meta-classes that were most similar with respect to physical watershed attributes (Figure 2-2). These seven groups were defined as the *major stream classes* (i.e. meta-classes) and carried forward for subsequent analyses. Classes 2 and 4 contain the most stream segments, accounting for approximately 55% of total linear distance mapped. Classes 3, 6, and 7 each contain less than 10% of the distance mapped. The global PERMANOVA of all 7 classes showed a significant difference between at least one class (pseudo-$F_{6,6993}$ = 775.12, $p<0.001$). Further pairwise comparisons showed that all 7 classes were significantly different from other classes (Table 2-4).

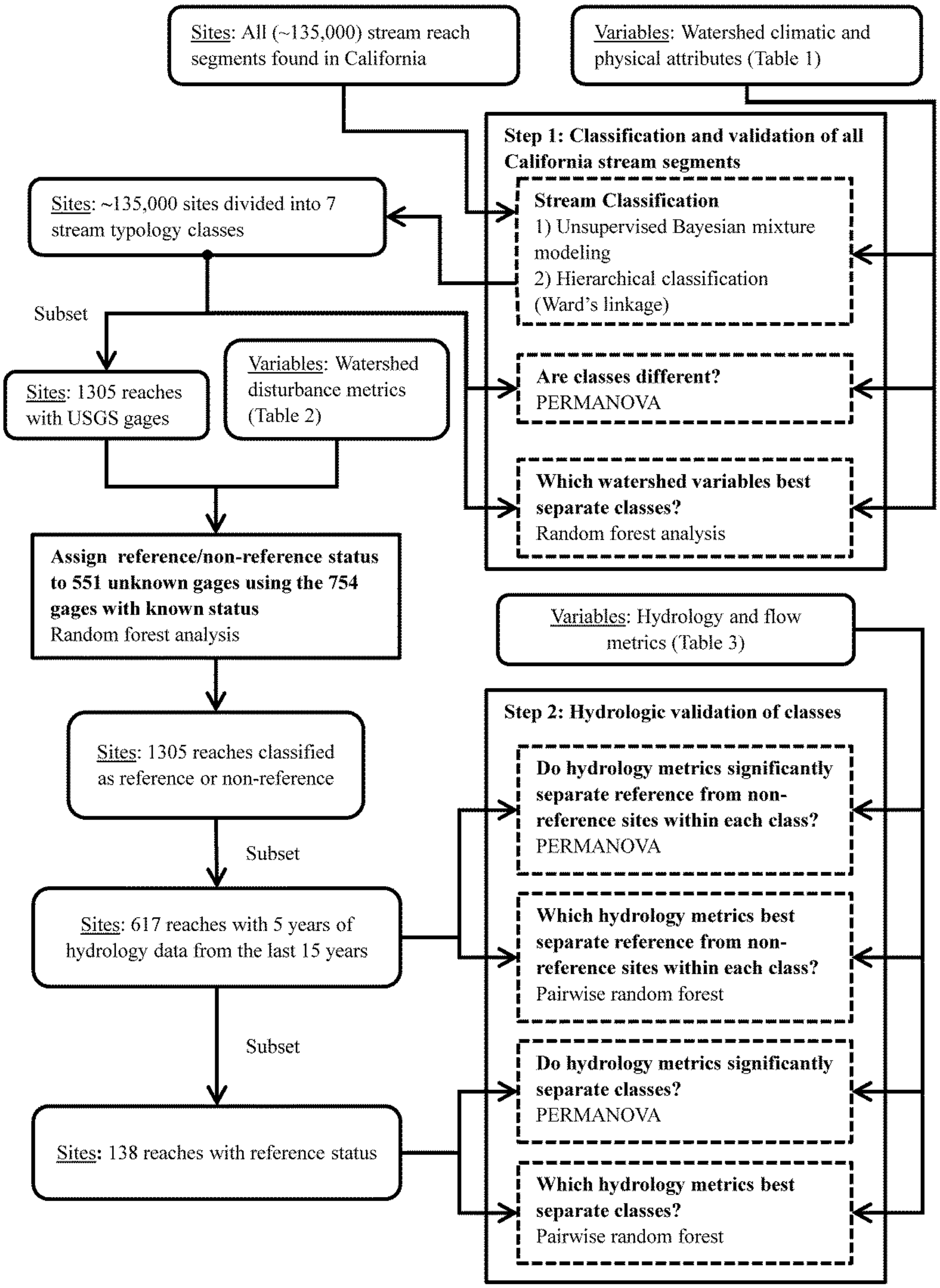

Figure 2-2. Cluster dendrogram showing aggregation of the 109 originally identified stream segment classes into seven proposed major stream meta-classes. The bar graphs between each split in the dendrogram show the three variables that the random forest analysis indicated as the most important in differentiating the two groups produced by the split. The values on both sides of each bar graph are each group's mean of the most important variable (with standard error in parentheses).

Random forest analysis separated the seven aggregated classes according to differences in winter precipitation, geology, soil characteristics, and mean watershed elevation (Table 2-5, Figure 2-2), with extremely low OOB misclassification rates of <4%. Class 1 streams (Figure 2-3a) are located in the drier mountainous regions of the state, but also found in the foothills of the southern Sierra Nevada Mountains and the mountains surrounding Scott Valley in northern California. Class 2 streams (Figure 2-3b) are mostly located in the Chaparral region, comprising streams located in the foothills surrounding the Central Valley and the hills and valleys west of the Central Valley, as well as in plains of the South Coast. Class 3 streams (Figure 2-3c) are consists of streams in the drier, volcanic Modoc Plateau in northeastern California and northern Sierra Nevada Mountains, although some large river segments of the Central Valley and southwestern deserts fall into this group. Class 4 streams (Figure 2-3d) are mostly distributed amongst the wetter North Coast, northern Sierra Nevada and northern Chaparral. Class 5 streams (Figure 2-3e) are almost exclusively located in dry southwestern desert. Class 6 streams (Figure 2-3f) are almost exclusively glaciated Sierra Nevada region. Finally, class 7 stream segments (Figure 2-3g) are located in the Central Valley and surrounding Chaparral foothills.

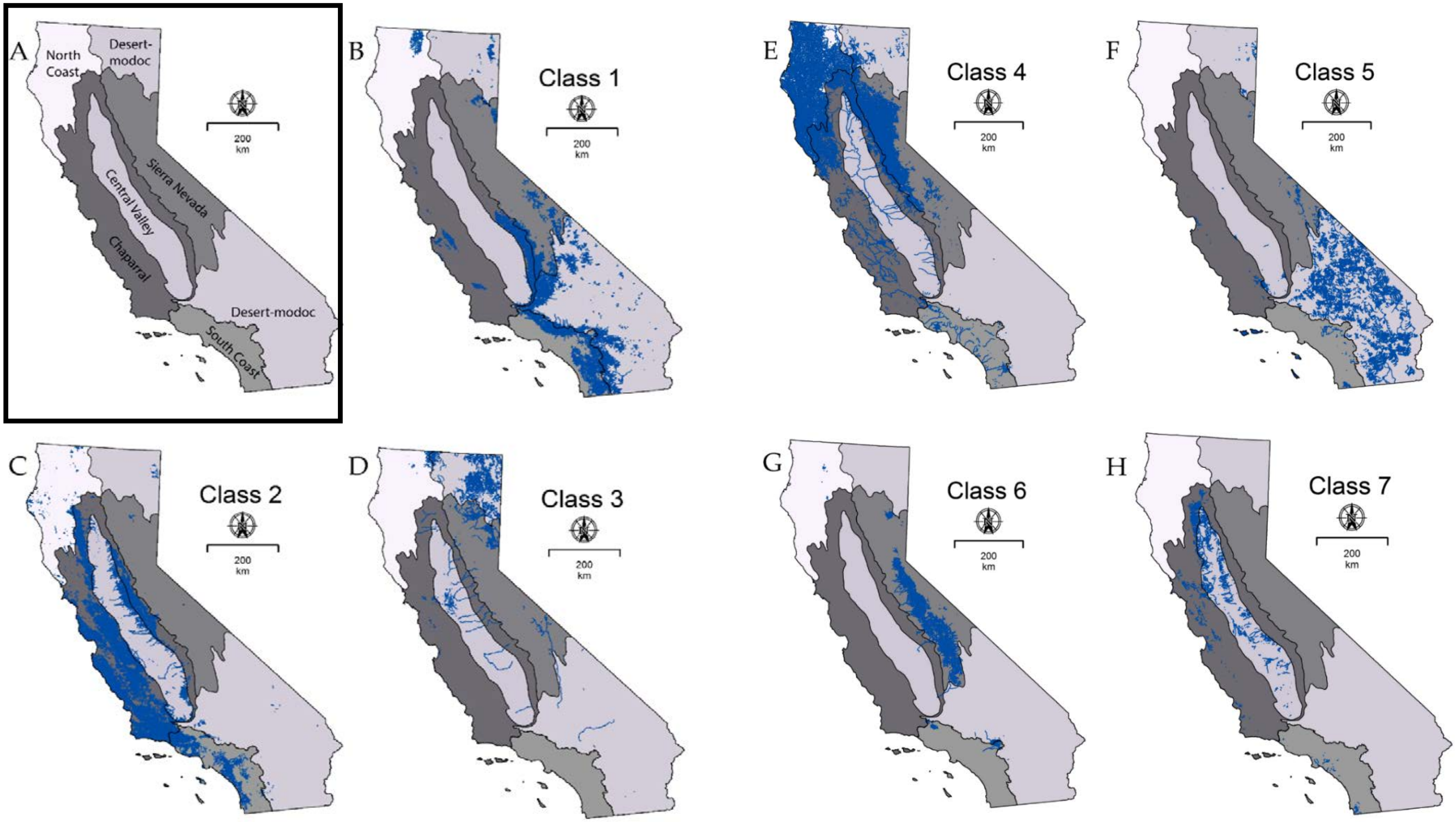

Figure 2-3. Maps showing the distribution of each stream class in the state. The Perennial Stream Assessment (PSA) regions are defined in panel A, with each of the 7 classes shown in panels B-I. Each stream segment assigned to each class is represented by the blue lines.

Table 2-4. The results of the pairwise PERMANOVAs of 7 classes using watershed climate and physical variables. One thousand stream reaches were randomly selected from each class for the analysis.

Class comparison			$\text{Pseudo-F}_{1,1998}$	p-value
One	vs	Two	451.81	<0.001
One	vs	Three	385.78	<0.001
One	vs	Four	427.46	<0.001
One	vs	Five	488.71	<0.001
One	vs	Six	783.44	<0.001
One	vs	Seven	1099.68	<0.001
Two	vs	Three	584.56	<0.001
Two	vs	Four	347.35	<0.001
Two	vs	Five	803.07	<0.001
Two	vs	Six	1630.07	<0.001
Two	vs	Seven	553.92	<0.001
Three	vs	Four	422.51	<0.001
Three	vs	Five	484.47	<0.001
Three	vs	Six	741.20	<0.001
Three	vs	Seven	682.48	<0.001
Four	vs	Five	942.68	<0.001
Four	vs	Six	774.61	<0.001
Four	vs	Seven	791.24	<0.001
Five	vs	Six	1220.81	<0.001
Five	vs	Seven	897.97	<0.001
Six	vs	Seven	2014.93	<0.001

Table 2-5. The general climatic, geologic, soil, and physiographic characteristics of each class and the proportion of each class located in each of California's Perennial Stream Assessment (PSA) regions. The regions are as follows: C = Chaparral, CV = Central Valley, SC = Southern Coast, S = Sierra Nevada, D-M = Desert-modoc, and NC = North Coast

Class	Winter precipitation	Geology	Elevation	Soils	C	CV	SC	SN	D-M	NC
One	Intermediate	No sedimentary		Moderate organic	0.16	0.01	0.27	0.13	0.4	0.03
Two	Intermediate	Sedimentary		Low thickness	0.73	0.13	0.12	0	0.01	0.01
Three	Intermediate	High volcanic	High	High organic	0.06	0.1	0	0.21	0.58	0.04
Four	Very high	Variable	Intermediate		0.25	0.03	0.02	0.25	0.03	0.41
Five	Low	Little sedimentary	Intermediate	Little organic	0.03	0.02	0.02	0.03	0.91	0
Six	High	No sedimentary Glaciated	Very high		0.01	0	0.04	0.9	0.04	0
Seven	Intermediate		Very low		0.32	0.65	0.02	0	0	0

2.3.2 Determination of reference vs non-reference conditions

The RF analysis had an OOB of 14.19% for both reference and non-reference classes, with 88 non-reference sites being misclassified as reference (out of 600) and 19 reference sites being classified as non-reference (out of 154). The most important variable in separating reference from non-reference sites was the total volume of reservoirs in the watershed, with a mean decrease of accuracy of 185.9, which was more than 4 times more important than next important variables: mean impervious cover (43.0), proportion of freshwater withdrawals (39.9), proportion of high intensity developed land use (34.4), proportion of cropland (34.3), and population density (29.8). Five variables, watershed area, proportion of shrublands, proportion of mines, proportion of NPDES polluters, and proportion of canals, had negative importance values, implying that their inclusion decreased model accuracy, and were removed from the analysis. Once we calibrated our RF model using known reference/non-reference sites, we classified gauges with unknown reference status as reference or non-reference, with 280 gauges designated as non-reference gauges and 271 designated as reference.

2.3.3 Hydrologic validation of class membership

The pairwise PERMANOVAs showed that hydrology data effectively separated most classes, with the exception of classes 3, 4, and 6, which were not significantly (i.e., p-value <0.05) distinguished from each other (Table 2-6). These three classes contain streams with the highest maximum, mean, and minimum flows in the state. Additionally, the hydrological comparison of class 1 to class 5 was marginally significant at p=0.065. The random forest analyses had relatively low misclassification error rates for most pairwise comparisons (< 18%), except for the non-significant comparisons and the class 1 vs. class 2 comparison (Table 2-6). Most of the hydrologic variables that best discriminated between-stream typology classes were indicative of high flows, mean or median flows, or flow timing (Table 2-6). The mostly desert class 5, not surprisingly, had lower maximum, mean or median flows in relation to watershed area than streams in all other classes. Class 1 also had lower maximum, mean or median flows than streams than classes 2, 3, 4, and 6. The chaparral hills and valley streams of class 2 had a greater occurrence of extremely low flows than class 3 and lower maximum or mid-flows than class 4. The mostly Sierra Nevada Class 6 differed from classes 1, 2, and 4 by the month of maximum flow, with the latter having maximum flows during early summer and the former having maximum flows during winter months, an indication of snowmelt streams in class 6 and the Mediterranean and humid streams in classes 1, 2, and 4. Class 6 streams were also less stable than classes 2 and 4. Although the relationships were not significant, the five reference streams in class 3 had greater periods of time between low and high flow events compared to the humid 4 and mountainous class 6.

Table 2-6. The results of the pairwise PERMANOVAs and random forests of 6 classes using hydrology data from reference condition streams. DF = degrees of freedom; F = pseudo-F; OOB = out-of-bag error rate (i.e., misclassification rate); MDoA = mean decrease of accuracy.

		PERMANOVA				Random forest			
Class comparison	Number of Sites	DF	F	p-value		OOB Error	Important variables	MDoA	Variable mean per class
One vs Two	26 vs 30	1, 50	4.04	0.009		23.00%	S_Qmax	43.83	0.11 vs 0.4
							S_Q99	32.26	0.1 vs 0.32
One vs Three	26 vs 5	1, 8	4.84	0.013		0.00%	QmeanIDR	22.12	0.4 vs 1.74
							Qmed	16.14	0.05 vs 3.13
							Q50	15.88	0.05 vs 3.03
One vs Four	26 vs 61	1, 50	12.75	0.001		17.00%	Q25	15.63	0.02 vs 1.82
							S_Qmean	30.84	0.01 vs 0.06
							S_Qmedian	27.59	<0.01 vs 0.06
One vs Five	26 vs 6	1, 10	3.59	0.065		16.00%	S_Q99	22.18	0.1 vs 0.6
							S_Qmax	25.33	0.11 vs <0.01
							S_Qmean	21.62	0.01 vs <0.01
One vs Six	26 vs 10	1, 18	7.47	0.002		11.00%	S_Q99	18.82	0.1 vs <0.01
							MaxMonth	73.04	2.81 vs 4.9
							S_Qmedian	14.31	<0.01 vs 0.04
Two vs Three	30 vs 5	1, 8	4.35	0.024		9.00%	S_Q95	13.43	0.03 vs 0.23
							Less1CFS	26.23	89.17 vs 16.38
							Q50	18.69	0.14 vs 3.03
							Qmed	18.67	0.14 vs 3.13

		PERMANOVA				Random forest			
Class comparison	Number of Sites	DF	F	p-value		OOB Error	Important variables	MDoA	Variable mean per class
Two vs Four	30 vs 61	1, 58	9.67	0.001		16.00%	S_Q75	41.96	0.01 vs 0.06
							S_MxMnthQ	36.65	0.03 vs 0.13
Two vs Five	30 vs 6	1, 10	4.94	0.023		6.00%	Q50	27.25	0.14 vs 2.91
							S_Qmean	20.46	0.02 vs <0.01
							S_Q99	19.16	0.32 vs <0.01
Two vs Six	30 vs 10	1, 18	5.34	0.005		3.00%	S_MxMnthQ	18.69	0.03 vs <0.01
							RBI	60.24	0.5 vs 0.14
							MaxMonth	59.53	2.3 vs 4.9
Three vs Four	5 vs 61	1, 8	2.09	0.206		11.00%	S_MxMnthQ	9.62	0.03 vs 0.16
							NoDisturb	26.28	219 vs 152.95
							QminIDR	10.94	0.94 vs 0.37
Three vs Five	5 vs 6	1, 8	5.86	0.009		0.00%	Less1CFS	10.6	16.38 vs 60.3
							S_MxMnthQ	10.68	0.04 vs <0.01
							Less1CFS	10.64	16.38 vs 99.85
Three vs Six	5 vs 10	1, 8	2.72	0.095		20.00%	MxMnthQ	10.55	11.02 vs 0.02
							S_MxMnthQ	16.32	0.04 vs 0.16
							NoDisturb	15.5	219 vs 162.65
							S_Q95	13.67	0.08 vs 0.23

		PERMANOVA				Random forest			
Class comparison	Number of Sites	DF	F	p-value		OOB Error	Important variables	MDoA	Variable mean per class
Four vs Five	61 vs 6	1, 10	5.72	0.007		4.00%	S_Qmean	14.09	0.06 vs <0.01
							S_MxMnthQ	14.02	0.13 vs <0.01
Four vs Six	61 vs 10	1, 18	1.7	0.189		20.00%	S_Q99	13.61	0.6 vs <0.01
							MaxMonth	51.19	2.8 vs 4.9
							RBI	22.51	0.33 vs 0.14
Five vs Six	6 vs 10	1, 10	6.79	0.007		13.00%	HighNum	15.89	5.52 vs 3.2
							S_Qmean	14.56	<0.01 vs 0.05
							S_Q90	14.53	<0.01 vs 0.15
							S_Q50	14.4	<0.01 vs 0.01

Hydrologic differences between reference and non-reference gauges were not always detectable (Table 2-7). Our results may not represent all of the hydrologic distinctions between classes because of uncertainty in the estimates of flow metrics for sites with relatively short periods of records that include unusually wet or dry years. Despite the reduction in explanatory power, the most important hydrologic variables differentiating at least some reference from non-reference streams were distinct from those differentiating reference streams among stream typology classes. Many of the variables that distinguish reference from non-reference sites within stream classes are associated with low flow conditions, flashiness, or streamflow recession rates, while maximum and mean or median flow was often most important in differentiating among the natural stream typology classes (Table 2-7). Storm-flow recession (SFR) is higher in non-reference sites in the chaparral-dominated classes 1 and 2, while non-reference sites have a greater number and intensity of high flows in class 1 and are flashier in class 2. Mean and median flows are important in class 3, with non-reference sites having greater flows. Disturbed gauges in class 4 are distinguished by extremes, i.e. a greater number of short duration low-flow events and lower high flows). The distinction between reference and non-reference gauges in class 5 is significant and dominated by low flows and the number of high flow events. Reference streams in class 5 have extremely long periods of low flows and relatively few high flow events compared to non-reference streams. Non-reference gauges in class 6 have lower base-flow recession rates, reduced daily changes in streamflow, reduced maximum flows, and greater durations of low flows. The variables highlighted from the random forest analysis are sensitive to anthropogenic alterations and provide some compelling insights to regional disturbance. The lack of significance in most of the PERMANOVAs indicated a need for further analysis in subsequent studies, including the exploration relationships of these hydrology metrics with instream biology.

Table 2-7. The results of the pairwise PERMANOVAs and random forests comparing reference condition reaches to non-reference condition reaches for 6 classes using hydrology data. DF = degrees of freedom; F = pseudo-F; OOB = out-of-bag error rate (i.e., misclassification rate); MDoA = mean decrease of accuracy.

			PERMANOVA				**Random forest**				
Class	**# ref**	**# non-**	**DF**	**F**	**p-value**		**OOB**	**Important**	**MDoA**	**Ref mean**	**Non-ref**
One	26	59	1, 50	1.64	0.287		24.00%	SFR	38.61	-1.6	-2.15
								Q99	35.66	2.56	11.44
								HighNum	31	3.13	4.58
Two	30	82	1, 58	1.33	0.332		23.00%	RBI	33.43	0.5	0.67
								SFR	33.22	-0.42	-1.56
								S_Q90	21.97	0.04	0.02
Three	5	52	1, 8	1.98	0.26		21.00%	QmeanIDR	16.79	7.1	38.05
								Q75	13.45	6.16	26.72
								Q95	12.82	19.95	85.53
Four	61	226	1, 120	2.25	0.199		17.00%	LowNum	43.37	1.4	2.55
								S_Q75	29.33	0.06	0.03
								LowDur	27.42	37.33	29.81
Five	6	4	1, 6	4.09	0.041		40.00%	HighNum	19.15	1.5	9.13
								LowDur	17.82	528.5	7.88
								MaxMonth	9.12	1.33	2.25
Six	10	56	1, 18	1.75	0.279		24.00%	BFR	31.35	-0.05	-0.02
								S_MxMnthQ	22.09	0.16	0.06
								PDC50	20.85	0.07	0.03

2.4 Discussion

Our use of a combined deductive-inductive approach to classify streams provides some distinct advantages over the single-track (i.e., deductive or inductive only) classifications. The inductive classification approach requires an extensive array of reference condition gauges spanning environmental gradients sufficient to capture the natural variation in streams (Olden et al. 2012). California streams, like most developed regions, are extensively hydrologically altered and regulated (Carlisle et al. In Review) and an inductive approach, relying solely on reference condition gauges, may bias the classifications towards small unregulated streams in mountainous or protected areas. A disadvantage of the deductive approach is rooted in the assumption that variation in stream hydrologic regimes is driven by the environmental variables used in the analysis (Olden et al. 2012). Although this is generally true, input data sets used to develop the deductive classification may not be available at a fine enough resolution to capture more localized factors that influence hydrologic regime at the sub-basin scale (Brown et al. 2014). Important hydrological variation in streams, such as perennialism versus intermittency, is not easily discerned using climate, geology, and geomorphology variables. A combined deductive-inductive approach allows us to provide universal stream classification regardless of availability of reference streams, while validating our classes using known hydrologic variables from reference gauges. We can then estimate the hydrologic differences between classes and assess how anthropogenic hydrologic alteration influences flow characteristics. Furthermore, the combined/hybrid approach allows us to estimate flow under “minimally disturbed conditions” based on the deductive analysis and compare to actual flow based on the inductive analysis in order to estimate hydrologic change.

Our classifications are similar to classifications developed of the United States using a deductive approach, classifying watersheds according to environmental characteristics. Wolock et al. (2004) classified all ~200 km^2 watersheds in the United States using environmental variables they selected *a priori* as important factors driving hydrological variation in streams and rivers: elevation, bedrock and soil permeability, and the amount of precipitation minus potential evapotranspiration. The most important variables driving our classifications, elevation, winter precipitation and proportion of sedimentary bedrock, are very similar to their selected variables, indicating our classification technique also highlights environmental drivers important in regulating hydrological variability. Wolock et al. classified most of California into semiarid mountains with impermeable bedrock (their classes 17 and 18), arid playas with permeable soils (class 14), arid plains with permeable soils (class 5), and humid mountains with permeable soils, but impermeable bedrock (class 16). The distribution of some of our classes corresponds fairly well with theirs, with our Class 6 lying within their semiarid mountain, impermeable bedrock, and permeable soil class, our Class 5 lying within their arid playa class, and our Class 7 lying almost within their arid plains class. However, our classification provided additional resolution, particularly among semi-arid regions in southern vs. northeastern California (classes 1 and 3) and different temperate regions (e.g. classes 2 and 4). This likely reflects the increased sensitivity of local scale models to subtle landscape properties, which are often aggregated for continental scale mapping.

Lane et al. (in press) recently classified California streams using a similar inductive-deductive approach, classifying according to streamflow data and validating using environmental variables.

They partitioned reference condition streams into 7 classes defined according to the primary drivers of hydrological variation in streams: snowmelt, rain, groundwater, ephemeral streams, and multiple transitional classes. While their classifications focused on various aspects of streamflow variation and ours focused on environmental drivers of hydrological variation, there was some overlap between classes. For example, stream reaches they defined as snowmelt were partitioned among our 3 classes with the highest elevations (1, 3, and 6), their *snowmelt and rain* gauges were found almost entirely within our Class 2 streams segments located along western edge of the Sierra Nevada Mountains, and their groundwater, transitional snowmelt, and rain-driven gauges mostly occur in our Class 4 stream segments. These relationships between the two classes highlight the differences between classifying stream segments on streamflow versus environmental setting. Their classes emphasized hydrological variation and seasonality while our divisions emphasize differences in the magnitude of streamflow, particularly the influence of high and low flow events. Both classifications can be helpful to stream managers in California, if managers' primary concern is the disruption of stream seasonality and variability or the magnitude of mean flow and/or high and low flow events. One stark difference between our classification and Lane et al.'s (in press) is the dearth of reference stream gauges found in our classes 3, 5, and 7. These three classes are either found in relatively arid or heavily agricultural regions. The lack of reliable, undisturbed streamflow data will result in an inadequate or inappropriate development of flow-based stream classes for these regions, while our environment-based classes provide such coverage. Further research is underway to reconcile both classifications based on the strengths and weaknesses of each classification.

Inductive approaches classify streams using flow data from gauges to create classes distinguished by hydrological variation (Poff and Ward; 1989, Poff, 1996; Kennard et al. 2010; Liermann et al. 2012). These approaches produce classes such as stable groundwater, snow-melt or rain driven. One distinct advantage of an inductive technique is the ability to distinguish intermittent streams from perennial streams. For example, Kennard et al.'s (2010) classification of streams in Australia resulted in 12 classifications, 8 of which were intermittent and were further classified by precipitation seasonality and consistency. Our classification technique does not inherently separate intermittent from perennial streams, but does differentiate climactic, geomorphological, and geological drivers that make a stream more likely to be intermittent, which may be important for future assessment of hydrologic alteration. Additionally, one inherent weakness associated with using reference gauges to classify streams may be the fact that some regions or stream types are under-represented in the analysis, particularly dry or largely agricultural areas (Liermann et al. 2012) or large rivers (Kennard et al. 2010), all present in California. Initial classification based on environmental variables ensures representation of all stream segments when reference conditions are not present (Olden et al. 2012).

The hybrid approach used in this study also reveals unique interactions among climate, physical setting and anthropogenic disturbance in shaping stream flow characteristics across California. Our analysis shows that stream typology is shaped mainly by patterns of large flow events as influenced by major features of the landscape such as, elevation, slope and geology. In contrast, deviation from reference is largely defined by changes in low flow variables, average daily flow, duration of flow, and timing of low and high flow events. This is consistent with observations that changes in land use have a substantially greater effect on annual peak flows compared to infrequent floods (Hollis 1975). For example, annual flow events (1 year recurrence interval) may change more than 10-fold with less than 5% increase in urbanization, whereas large storm events may only double with more than 30% increase in urbanization (Hollis 1975).

Comprehensive stream classification provides the foundation for establishing desired flow regimes necessary to support biological communities and ecological functions of management importance (Poff et al. 1997). Application of flow-ecology management using the ELOHA framework begins with stream classification. This is particularly important in places like California which have extreme climatic, altitudinal, and geologic gradients which affect stream morphology and flow patterns. This complexity, combined with spatially variable patterns of land use (e.g. urban, agricultural, timber, hydropower) produces highly variable flow responses in streams that must be accounted for during development of flow management targets. For example, disturbance of dry mountainous streams with a propensity for low flows (Class 1), will rapidly exacerbate high flow events, and subsequent geomorphic alternation such as avulsion or incision. In contrast, anthropogenic changes in catchments of non-sedimentary, semi-arid low mountain streams (Class 3) tend to change storage properties (based on high soil permeability), which likely makes these stream susceptible to changes in mean and median flows. The importance of climatic gradients to stream classification is seen in the non-sedimentary, arid streams (Class 5). These desert streams are naturally dry and have been shown to respond with dramatic increases in intensity and duration of flow following land use alteration (Schriever et al 2015). Finally, the natural flow variability in wet, humid high mountain streams (Class 6) tend to be muted indicating the importance of the water withdrawals and dams associated with disturbed streams in the Sierra Nevada Mountains. The time scale of our variables may not be able to address sub-daily alteration of flow associated with hydropower, a main disturbance in the region

The comprehensive classification derived from the deductive portion of our analysis will allow application of the ELOHA analysis throughout the entire state of California. By not relying solely on an inductive classification, we avoid the limitation of areas with few or no reference condition gauges available, leading to classes that are biased towards small streams in uninhabited areas. Our deductive-inductive approach provides an alternative classification technique that creates hydrologically discriminant classes for all stream segments. This will allow us to partition the statewide analysis into relatively homogenous subgroups thereby increasing the ability to develop meaningful relationships between hydrologic variables and biological response metrics. The inductive portion of the analysis provides important insight into which of the hundreds of available hydrologic variables are most likely to produce meaningful flow-ecology relationships. For example, low flow and flow duration variables were important determinants of reference condition in several classes. Past work in Mediterranean and dry climates has shown that changes in the duration of wet vs. dry periods and flow intermittence are key determinants of invertebrate community composition (Datry 2012). Ultimately, if these variables produce strong explanatory relationships with biological variables that are also indicative of deviation from reference, they will be good candidates for development of flow targets for future watershed management. Classification allows such relationships to be "tuned" to regional flow patterns and to focus on the aspects of flow-modification that are most critical for each specific stream class.

3.0 ASSESSMENT OF HYDROLOGIC ALTERATION

3.1 Background

Assessing hydrologic alteration at all sites of biological interest is critical to establishing flow-ecology relationships. Unfortunately, relying on empirical flow data from available gauges to understand the regional extent of alteration or develop regional flow-ecology relationships presents a serious constraint as long-term flow data is often limited (Puckridge et al. 1998; Poff et al. 2006). Additionally, the extent of hydrologic alteration should be based on a comparison of contemporary conditions to reference conditions prior to watershed development. This is also extremely hard to quantify since flow records that date back to pre-disturbance period are rare at gauged sites and nonexistent at ungauged sites (Carlisle et al 2010).

Modeling provides an alternative for estimating both current and reference streamflows at gauged and ungauged locations as a foundation for developing regional flow-ecology relationships (Poff 2009). Typical approaches for predicting streamflow in ungauged basins are based on transferring gauged data at drainage basin scales (Sivapalan et al. 2003). This transfer is usually done by either establishing regression relationships between the different flow metrics or components of the hydrograph and the basin characteristics, or by estimating model parameters values from the gauged basin and inputting into hydrologic models applied to ungauged sites (Post and Jakeman 1999: Sivapalan et al. 2003: Wagener and Wheater 2006; Sanborn and Bledsoe 2006; Yadav et al. 2007; Wagener and Montanari 2011; Parajka et al. 2013; Buchanan et al. 2013). Reference condition is estimated by developing models at gauges in relatively unaltered settings in a space for time substitution. This concept of regionalization or using gauged basin behavior to predict flows at ungauged basins has been explored in many studies, but with limited success (Kokkonen et al. 2003; Moretti and Montanari 2008; Samaniego et al. 2010) given that flow regimes are inherently variable, dictated by geography, climatic patterns, and catchment properties (Poff and Zimmerman 2010). Though regression models and hydrological model performance can be comparable in terms of predicting daily flows, and flow metrics, regression models can be restrictive for exploring management scenarios, such as evaluating the impact of stormwater capture on streamflow in a watershed. Complex hydrological models can require significant effort during calibration and do not transfer easily to other sites at regional scales (Sivapalan et al., 2003). Simple yet representative hydrological models can be easily transferred to any stream reach of interest and, therefore, provide a viable approach for developing flow data necessary for regional flow-ecology relationships.

Our goal is to estimate hydrologic alteration at ungauged stream reaches in Southern California where we have bioassessment data. We developed an ensemble of simple, yet regionally representative hydrological models that can be easily transferred to any ungauged location in the region. We evaluate the ability of these models to estimate current and reference condition flows at ungauged stream reaches, estimate biologically relevant flow metrics and provide a regional understanding of hydrological alteration in southern California.

3.2 Methods

Regional hydrological alteration was estimated using an ensemble of models calibrated and validated to represent watershed conditions in the southern California study area. Models were developed for 32 gauged catchments with sufficient data availability and that represented the range of physical watershed characteristics observed in the region. Calibration was based on biologically relevant flow properties in addition to traditional hydrograph fitting. Random forest modeling was used to assign one of the 32 calibrated and validated models to 799 ungauged bioassessment sites in the study area. Current and historical flows were modeled for each bioassessment site using a standard precipitation time series. The difference in flow was used to estimate the extent of hydrologic alteration across the region using a selected set of commonly used flow metrics.

3.2.1 Hydrology, precipitation, and GIS data

We selected 32 USGS gauges with available hourly flow records and hourly precipitation data for development of the model ensemble (Figure 3-1a). The flow and precipitation data at these gauges overlapped for at least a three-year period representing dry, wet and average years in California. These gauges were selected to represent a range of watershed conditions (for example, imperviousness, landuse, groundwater storage) found in Southern California. Basic catchment properties for each station were compiled in GIS for use in model parameterization. Hourly precipitation data was sourced from national databases (Figure 3-1b): National Oceanic and Atmospheric Administration (NOAA), National Weather Service Automated Local Evaluation in Real Time (ALERT), state database: California Irrigation Management Information System (CIMIS) and California Data Exchange Center (CDEC) and local databasees: San Diego Regional (SDRCD), Ventura County Watershed Protection District (VCWPD).

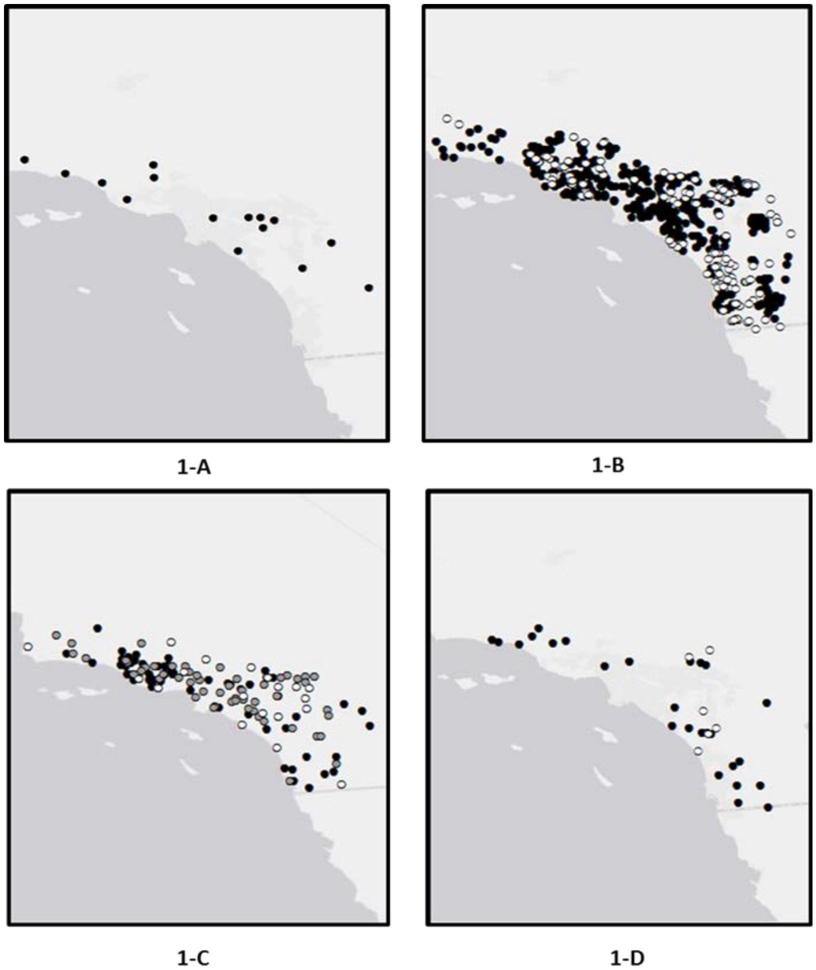

Figure 3-1. Four panel map showing 32 gauged catchments for model ensemble (1-A, black dots for 26 gauges finally selected), 296 precipitation gauges used in the study (1-B, the white dots are cutoff at 1% and grey dots at 45%), 799 'ungauged' bioassessment sites (1-C, black dots where we were able to predict precipitation), 15 'ungauged sites with flow data used for validation (1-D).

3.2.2 Model Calibration

We used the HEC-HMS rainfall-runoff modeling platform to simulate flows at the ungauged sites. HEC-HMS can represent most of the critical hydrologic processes of a watershed system with a parsimonious set of model parameters, making it a good choice for development of the regional model ensemble. We chose HEC-HMS over more detailed models such as Hydrological Simulation Program-Fortran (HSPF) or Gridded Surface Subsurface Hydrologic Analysis (GSSHA) to avoid the need for an intensive calibration process involving a large parameter set when applying to high numbers of ungauged basins (Sivapalan et al 2003). The HEC-HMS models were developed for 2005, 2006, and 2007, representing wet, normal and dry years respectively. The models were parameterized to account for infiltration losses, transformation of excess precipitation to runoff, and baseflow contribution to subbasin outflow using the input data shown in Table 3-1.

Table 3-1: Parameters used to develop HEC-HMS models for application to the ungauged sites. Parameters in bold were measured or estimated, remaining parameters were calibrated. In case of application to ungauged site, the observed flow column is left empty.

	Parameters
Measured or Estimated	**Area**
	Imperviousness
	Observed flow
	Observed precipitation
	Time of concentration
	Initial Loss
Calibrated	Maximum Storage (in)
	Initial Storage (%)
	Maximum Storage (in)
	Initial Storage (%)
	Initial Deficit (in)
	Maximum Deficit (in)
	Constant Rate (in/hr)
	Ground Water (GW) 1 Initial Discharge (cfs)
	GW 1 Storage Coefficient (hr)
	Number of GW 1 Reservoirs
	GW 2 Initial Discharge (cfs)
	GW 2 Storage Coefficient (in)
	Number of GW 2 Reservoirs

Drainage area of the streamflow gauges was delineated in ArcGIS 10.1 using National Elevation Dataset (NED) 10 m DEM (Gesch et al. 2002). Total imperviousness was computed for each basin by clipping 2006 NLCD data (Fry et al. 2011) in ArcGIS. Time of concentration (TOC) was calculated using the Kirpich Method (Kirpich, 1940) and data obtained from the 10 m DEM ArcGIS delineations. The Clark Unit Hydrograph storage coefficient was calculated using Equation 1):

$$R = 0.37 * TOC^{1.11} * L^{0.80} * A^{-0.57} \quad [3\text{-}1]$$

Where *R* is the storage coefficient in hours, *TOC* is the time of concentration in hours, *L* is the channel flow length in mi and *A* is the basin area. Initial losses were estimated using a Soil Conservation Service (SCS) curve number to reflect losses associated with different landuses in the watershed. The initial loss in HEC-HMS is estimated as an initial abstraction of 0.2S, where S is the area weighted ultimate soil storage potential based on the composite SCS Curve Number in that watershed. Simple canopy, simple surface, and constant loss methods were used to simulate infiltration losses, while the linear reservoir method with two layers was used to represent baseflow contributions.

The HEC-HMS models were calibrated sequentially for four separate criteria. Two criteria emphasized overall fit: 1) Visual hydrograph matching and 2) Nash-Sutcliffe overall efficiency (NSE), and two emphasized metrics with relevance to instream biological communities (Gasith and Resh, 1999, Konrad et al. 2005, Morley and Karr 2002): 3) Extremely low flow periods indicative of stream drying (< 1cfs) and 4) Richard Baker Index (RBI) for flashiness (Konrad et al. 2008). Visual comparison was used as a baseline for further overall fit calibration using NSE, which is a measure of best overall fit and determines the accuracy of each model relative to the observed mean flow. Because it is calculated with respect to mean flow, NSE tends to be biased towards high flows (Jain and Sudheer, 2008) therefore tuning models for NSE alone may not accurately model streamflow flashiness and drying which are known to strongly influence stream biota in the region. To address this bias, we added a low-flow percent error (LFE) calibration to minimize the percent errors of time with flow less than 0.03 cms and a RBI percent error calibration to minimize the percent error between the observed and the predicted flashiness as measured by the index. Three separate calibrated parameter sets unique to each model for the NSE, LFE, and RBI criteria were compared, and a final set of optimal models was produced based on overall performance of these three quantitative calibration criteria. Table 3.2 lists the 26 gauged stations.

Table 3-2. List of sites used for the final ensemble, associated characteristics (size, imperviousness, and elevation) and flow and precipitation gauges used to calibrate the HEC-HMS models.

Site Name	Size (mi2)	Impervious	Elevation (ft)	USGS Gauge	Precipitation Gauges
Andreas	8.65	0	800	10259000	CDEC, CIMIS, NOAA
Arroyo Seco	16	0.46	1398	11098000	CDEC , CIMIS, NOAA
Arroyo Trabuco	54.12	19.06	80	11047300	CDEC, CIMIS, NOAA
Campo	84.11	0.55	2179	11012500	CDEC, SDCFCD
Carpinteria	13.1	0.1	130	11119500	VCWPD 254
DeLuz	33	0.32	270	11044800	ALERT , CDEC, CIMIS , NOAA
Sweetwater	45.4	0.28	3269	11015000	CDEC, NOAA, SDCFCD
Devil Canyon	5.49	0.74	2080	11063680	CIMIS
East Twin	8.8	0.64	1590	11058500	CIMIS
Jamul	70.11	0.54	512	11014000	CDEC, CIMIS, SDCFCD
Los Angeles	158	27.34	663	11092450	CDEC , CIMIS , NOAA, VCWPD
Los Coches	12.17	9.39	560	11022200	ALERT, CDEC
Poway	42.44	20.66	300	11023340	CIMIS SDCFCD
Lytle	46.6	0.33	2380	11062000	CDEC
Matilija	47.8	0.01	1380	11114495	NOAA, VCWPD
Mission	8.38	4.77	140	11119750	CIMIS
Rainbow	10.21	3.7	500	11044250	ALERT, CDEC, CIMIS
San Jose	5.51	0.4	96	11120500	CIMIS
San Mateo	80.8	0.13	405	11046300	ALERT , CDEC, CIMIS, NOAA
Santa Maria	57.6	2.52	1294	11028500	CIMIS
Santa Paula	38.4	0.14	619	11113500	VCWPD
Santa Ysabel	111.43	0.1	848	11025500	CDEC CIMIS , NOAA, SDCFCD
Santiago	12.5	0.21	1340	11075800	CIMIS
Sespe Fillmore	252	0.05	565	11113000	VCWPD
Sespe Wheeler	49.5	0.09	3500	11111500	VCWPD
Sandia	19.67	1.27	380	11044350	ALERT, CDEC , CIMIS , NOAA

Temporal validation was conducted at 10 of the gauged watersheds out of the ensemble of 26 watersheds using data from periods outside of the three-year calibration run. No model parameters were changed during temporal validation, only precipitation data from the validation period was input into the model. To test the transferability of the models to other sites (spatial validation), a resampling technique (Jack-knifing) that leaves out the 'jackknife estimator' or in our case one gauged site from the dataset and uses the remaining (N-1) sample size to predict values was applied. The models were adjusted for same four parameters described in the methods section from the 'jackknife estimator' site and simulated over the three-year period (2005-2007).

3.2.3 Assigning a novel site to a gauged HEC-HMS model

We used a large bioassessment data set, consisting of 799 sites (Figure 3-1c) sampled under a variety of regional surveys to develop regional estimates of extent of hydrologic alteration. We focused on bioassessment sites as a precursor to a companion study aimed to establish biologically relevant regional flow targets (Mazor et al, in review).

To assign a HEC-HMS model to the 799 ungauged sites, calibration gauges were first clustered using a flexible beta, which is a hierarchical clustering method based on hydrologic similarity. A suite of 32 metrics (bold in Table 3-3) that represent different components of the hydrographs were estimated for the gauged sites using custom scripts in R (C. Konrad, personal communication). The metrics were rank-transformed to improve comparability of metrics across different scales. A principal components analysis of the rank-transformed metrics was used to combine redundant metrics into independent synthetic gradients using the *prcomp* function in R (The R Core Team 2016). The first eight components (i.e., the number required to capture 95% of the variance in the flow metrics) were used as the basis for clustering. Hydrologic dissimilarity among gauges was calculated as standardized Euclidean distance using the daisy function in the cluster package in R (Maechler et at. 2015). A dendogram was created using flexible-beta (beta: -0.25) based on this dissimilarity matrix to visually define groups of gauges with similar flow metrics.

We developed a random forest model to assign cluster membership for calibration gauges based on watershed characteristics measured in GIS. Twenty-seven candidate predictor variables were evaluated for this model, representing both natural (e.g., climate, geology, elevation) and anthropogenic (e.g., road density, percent impervious) gradients. Recursive feature elimination (RFE function, Caret package, Kuhn et al 2012) was used to select the variables that were most useful in predicting cluster membership. RFE attempts to find the simplest model whose accuracy was with 1% of the most accurate model. The selected variables were used to calibrate the final 1000-tree assignment model using the random Forest package in R (Liaw and Wiener 2002).

To assign a HEC-HMS model to an ungauged test site, the random forest model was run with a single test site and calibration site data simultaneously. Proximity was then calculated as the frequency that the test site was assigned to the same group as a calibration gauge. The HEC-HMS model based on the most proximal gauge was then assigned to the ungauged site. This assignment was repeated for each of the 799 bioassessment sites.

3.2.4 Estimating hourly precipitation at the ungauged sites

A primary input to the HEC-HMS models at the ungauged site is hourly precipitation data for the period selected for estimating flows. Precipitation gauges are limited in the region, therefore we used an inverse distance weighting (IDW) interpolation method to predict hourly precipitation at the 799 sites using measured precipitation data from 206 precipitation gauge network in Southern California. For each ungauged location (P_a), we used a network of N (206) gauge stations P_i ($i = 1,\ldots N$):

$$P_a = \frac{\sum_{i=1}^{N} w_{a,x} \cdot P_i}{\sum_{i=1}^{N} w_{a,i}} \qquad \textbf{Eqn 3-2}$$

where $w_{a,i} = \frac{1}{(D_{a,i})^p}$ is a weighting function, *a* denotes an interpolated (assessment) point, *i* is an interpolating (known) point, *D* is the distance (metric operator) from the known point *i* to the unknown point *a*, N is the total number of known points used in interpolation.

The IDW method was validated by randomly holding back 10 precipitation gauges and predicting precipitation by interpolating the rest of the dataset. This was repeated 20 times, and in case of repeat predictions, the results were averaged.

For each of the 799 sites, the IDW prediction accuracy was tested by comparing the predicted annual aggregated values to observed aggregated values sourced from each county. Anomalously large or small values of precipitation, usually resulting from bad measured data used in the model were eliminated by setting upper and lower bounds using the measured maximum and minimum precipitation values (rounded to the nearest inch/year) by elevation for each county. For each year where modeled precipitation fell outside the multi-county measured range the site was removed from the analysis. This process yielded 572 sites, where we could reliably predict precipitation.

3.2.5 Estimating flows under current and reference conditions, and hydrologic alterations

The HEC-HMS models matched to the ungauged sites were simulated to predict continuous hourly flows for 1990-2013 (current conditions) at each of the 572 sites. This 23-year period overlaps with the biological data collected at these sites. At each of the 572 ungauged sites the assigned HEC-HMS model was adjusted to each site by inputting site specific basin area, imperviousness, time of concentration, Clark Unit Hydrograph storage coefficient, and hourly precipitation data. We then selected a subset of 6 years that include two wet, two dry, and two average precipitation years for these sites. Dry years were defined as below the 30th percentile, average years between the 30th and 60th percentiles and wet years exceeding the 60th percentile of the total annual precipitation. In cases where more than two quality years existed, two were selected randomly from within the category. Model performance was validated using the limited measured flow data available at a subset of the 572 sites (N=15). Validation at each of the 15 sites was done for multiple years generating 67 combinations of sites and years.

Reference condition at ungauged sites were simulated by adjusting the current models to reflect pre-urbanization conditions: by setting imperviousness to zero to mimic no urban landuse, and by increasing initial losses to account for greater land availability. The same precipitation data used to estimate current flows were used to estimate historic flows to ensure compatibility during estimation of delta hydrological metrics.

A suite of 39 flow metrics was calculated for 4 climate regimes (average rainfall, wet, dry, and overall, the three IDR metrics were estimated for just the overall conditions). Average, wet and dry are based on precipitation conditions for 2 years each, whereas the overall metrics are estimated based on 6 years. This produced a total of 147 flow-precipitation condition combinations used to estimate hydrologic alteration (as the difference in metric values between current and historic conditions). The flow metrics are grouped into duration (n= 5), frequency (n= 4), magnitude (n= 16), timing (n= 7), and variability (n=7).

Estimated values of flow metrics were validated by correlating predicted values against values observed at 41 sites, including 26 gauged sites used for HEC-HMS ensemble and 15 validation gauges. Coefficient of determination $r^2>0.25$ were considered acceptable performance for the metric validation. Any metric with r^2 values lower than 0.25 were not considered acceptable, and were eliminated due to poor performance.

Hydrologic alteration was characterized as the difference in metric values between current and historic conditions. Magnitude metrics were further normalized by dividing by the historic metric value (or 1cfs or 0.03 cms, whichever was larger). Estimation of hydrologic alteration for duration and frequency flow metrics is complicated by the fact these metrics are based on comparison with benchmark high or low flow event (e.g. low flows are identified as flow below the 10th percentile), and this benchmark can change between historic and current conditions (e.g. the 10th percentile may decrease as a watershed undergoes urbanization). To deal with this changing baseline effect we applied an alternative approach of using the benchmark derived for historic conditions to calculate hydrologic alterations for duration and frequency metrics rather than independently estimating the metric value for historic vs current conditions, metrics for which we used this alternative approach are marked with an asterisk in Table 3-3).

Sites sampled as part of an ongoing regional ambient monitoring program that uses a probabilistic survey design were used to estimate the extent of hydrologically altered streams. The extent of regional alteration was estimated as the percent of stream-kms exhibiting an increase, decrease, or no change in the metric values. Extents were estimated for three land use classes (agricultural, undeveloped and urban), and three of the common stream classes in the regional data set (1, 2 and 4, Pyne et al. 2017). The extent of alteration (increasing or decreasing category) was estimated based on the percent of the stream kms assigned with no associated threshold. This contrasts with the alteration presented in Mazor et al. (this issue), which assesses alteration based on biological thresholds. Because sites were sampled under multiple surveys, weights were recalculated through post-stratification. These weights were used to estimate extent and magnitude using the Horvitz-Thompson estimator (Horvitz-Thompson 1952). Confidence intervals were based on local neighborhood variance estimators (Stevens and Olsen 2004). All calculations were conducted using the spsurvey package (Kincaid et al. 2013) in R (R Core Team 2012). Additional details about weight adjustments, land use classifications, and extent estimates are provided in Mazor (2015).

Table 3-3. Flow metrics (N=39) by the categories. Precipitation conditions are indicated by columns marked O (overall), W (wet), A (average), and D (dry). NA: Metric-precipitation condition combination not analyzed. Bold metrics (32) used for clustering. * Metric calculated using benchmarks derived from estimates of reference conditions

	Metric	Unit	O	W	A	D
Duration	HighDur*	days/event				
	Hydroperiod	proportion				
	LowDur*	days/event				
	NoDisturb*	days				
	Per_LowFlow	proportion				
Frequency	FracYearsNoFlow	proportion				
	HighNum*	events/year				
	LowNum*	events/year				
	MedianNoFlowDays	days/year				
Magnitude	MaxMonthQ	cms				
	MinMonthQ	cms				
	Q01	cms				
	Q05	cms				
	Q10	cms				
	Q25	cms				
	Q50	cms				
	Q75	cms				
	Q90	cms				
	Q95	cms				
	Q99	cms				
	Qmax	cms				
	Qmean	cms				
	QmeanMEDIAN	cms				
	Qmed	cms				
	Qmin	cms				
Timing	C_C	ratio				
	C_CP	ratio				
	C_M	ratio				
	C_MP	ratio				
	C_P	ratio				
	MaxMonth	month				
	MinMonth	month				
Variability	QmaxIDR	Unitless		NA	NA	NA
	QmeanIDR	proportion		NA	NA	NA
	QminIDR	cms		NA	NA	NA
	RBI					
	PDC50					
	BFR					
	SFR					

3.3 Results

3.3.1 Performance of model ensemble

The optimized ensemble of 26 models have NSE values ranging from 0.40-0.95, with an average value of 0.67 (Table 3-4). The NSE values for 22 of the 26 models are higher than 0.50. Percent LFE ranges from 0-59.50%, with a low value for this calibration criteria indicating good performance. The average value of the LFE is around 10%, and approximately 24 of the 26 models have less than 25% error. The RBI percent error is in the range of 0.1-55.3%, with 18 of the 26 models registering less than 25% error. Some of the best performing models are Los Angeles, DeLuz, Poway, and San Jose. Campo, and Sespe Wheeler Springs were the worst performing models.

Table 3-4. Model performance for the ensemble by sequential calibration (all values are of cumulative performance). Site names list the final set of 26 model sites, the remaining three column show sequential calibration for 1) Nash Sutcliffe Efficiency (NSE), Percent Low Flow Error, and Percent Richard Baker Index Error.

Site Name	NSE	Percent Error LF	Percent Error RBI
Andreas	0.58	5.3	8.1
Arroyo Seco	0.42	19.4	5.2
Arroyo Trabuco	0.73	15.7	33.2
Campo	0.49	7.4	55.3
Carpinteria	0.83	0.3	0.1
DeLuz	0.9	6.4	8.3
Sweetwater	0.57	2.9	71
Devil Canyon	0.57	59.5	27.8
East Twin	0.4	33.4	4.9
Jamul	0.46	15.8	13.8
Los Angeles	0.95	0	23.1
Los Coches	0.79	4.5	22.2
Poway	0.91	18.8	20.7
Lytle	0.44	18.9	6.3
Matilija	0.84	16.7	18.8
Mission	0.83	8	26.6
Rainbow	0.73	10.7	8.3
San Jose	0.81	5.5	25
San Mateo	0.75	2.5	50.7
Santa Maria	0.73	1.3	23.4
Santa Paula	0.53	0	17.8
Santa Ysabel	0.72	3.4	16
Santiago	0.58	0.1	27.6
Sespe Fillmore	0.61	0	22.6
Sespe Wheeler Springs	0.58	5.2	47.2
Sandia	0.63	0	24

3.3.2 Spatial, and temporal validation at gauged sites.

Temporal validation was conducted at 10 of the 26 gauges and results varied among them (Figure 3-2). Sites were validated for 2007-2010, which has two drought years, and one very wet year. At four of the gauges (Los Angeles, Arroyo Trabuco, Sandia, and Santa Paula), the performance remains comparable to the calibration period. However, the validation is poor at three sites (Campo, Santiago, and Arroyo Seco). Topography, and climatic factors, such as precipitation, strongly control the model performance. For example, Los Angeles, and Arroyo Trabuco are coastal watersheds with regulated flows, and little orographic control on precipitation. These models validate well, whereas, Arroyo Seco, located at the base of the San Gabriel Mountains with no flow regulations, steep terrain, and orographic control, validates poorly.

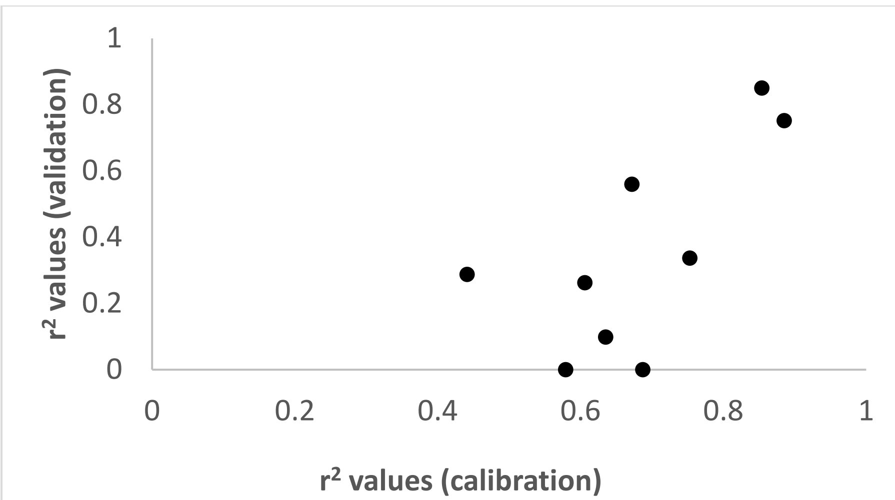

Figure 3-2. Temporal validation for 10 gauged models out of the ensemble of 26 models. The models were calibrated for 2005-2007 and the validation period varies for the sites. Typically for most sites data from 2007-2010 was used for validation.

Spatial validation using jack-knifing for the 26 gauged HEC-HMS models showed that 75% of the sites had a matched model that predicted flows at the site with NSE value >0.50 (Figure 3-3). The average NSE for validation was only 0.06 less than for calibration when the three poorly performing gauges (Campo, Lytle, and Devil Canyon) were excluded. Certain model parameters transferred better and produced the high NSE values for the ungauged sites. Models that calibrate poorly (for example, Campo: NSE =0.49, LFE=7.4, RBI =55.3; Lytle: NSE =0.44, LFE=18.9, RBI =6.3) also perform poorly during the jack-knifing (i.e. Campo parameters did not transfer well to other sites, and the performance at Campo using other model parameters remains low).

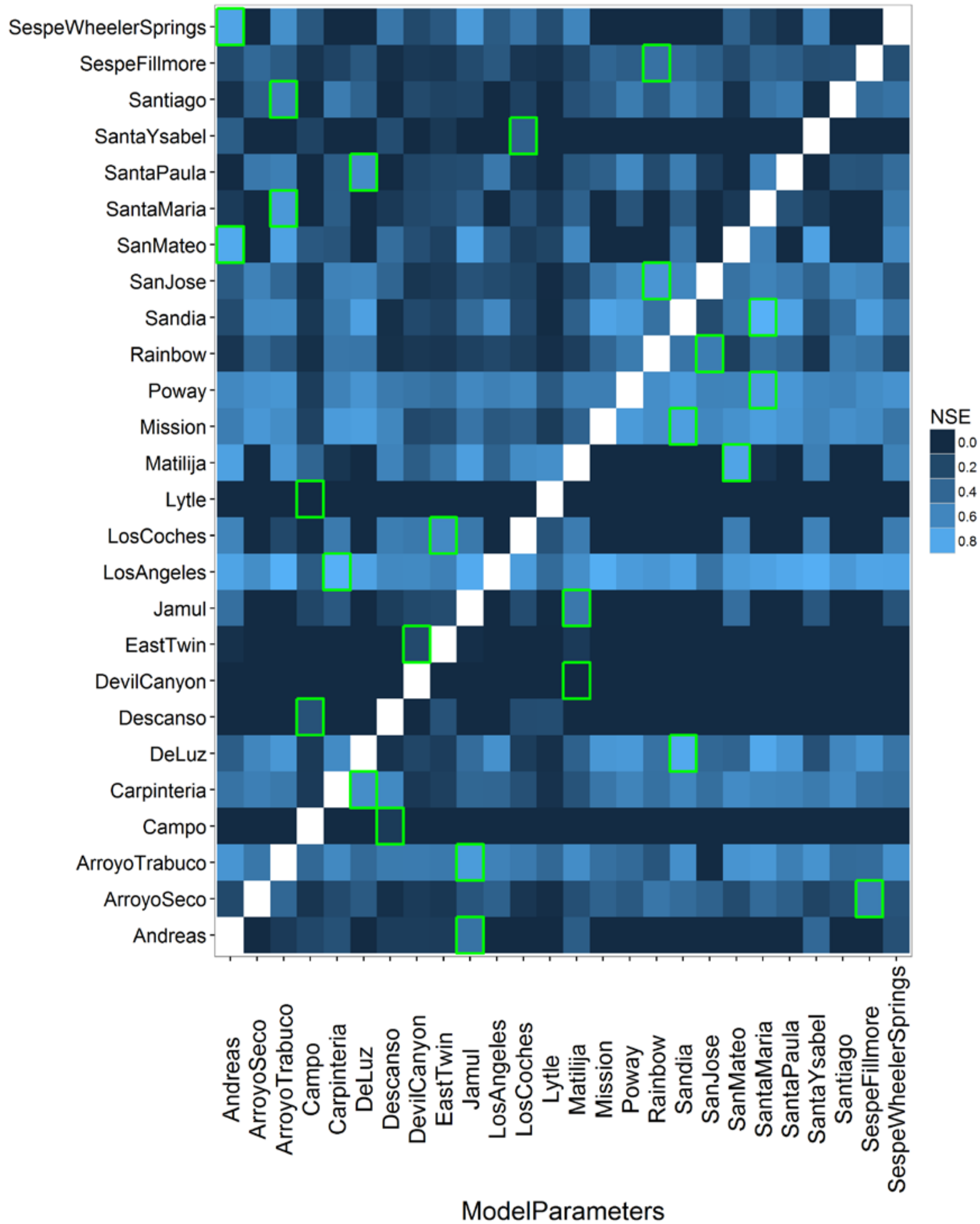

Figure 3-3. Jack-knife validation for 26 sites, y-axis shows the site treated as ungauged, and x-axis is for model parameters. Each row shows model performance fitted to 25 model parameters for a given ungauged site with green lined boxes highlighting the best performance based on NSE.

3.3.3 Cluster analysis and model assignment

Cluster analysis of calibration gauges yielded 8 groups, ranging in size from 3 to 5 gauges each. Generally, these groups did not show strong geographic clustering, as even the small groups included gauges that were spatially dispersed. Of the 29 candidate predictor variables, seven ranked highest in predicting group membership, with all but one (i.e., soil erodibility) variable representing anthropogenic factors (Figure 3-4).

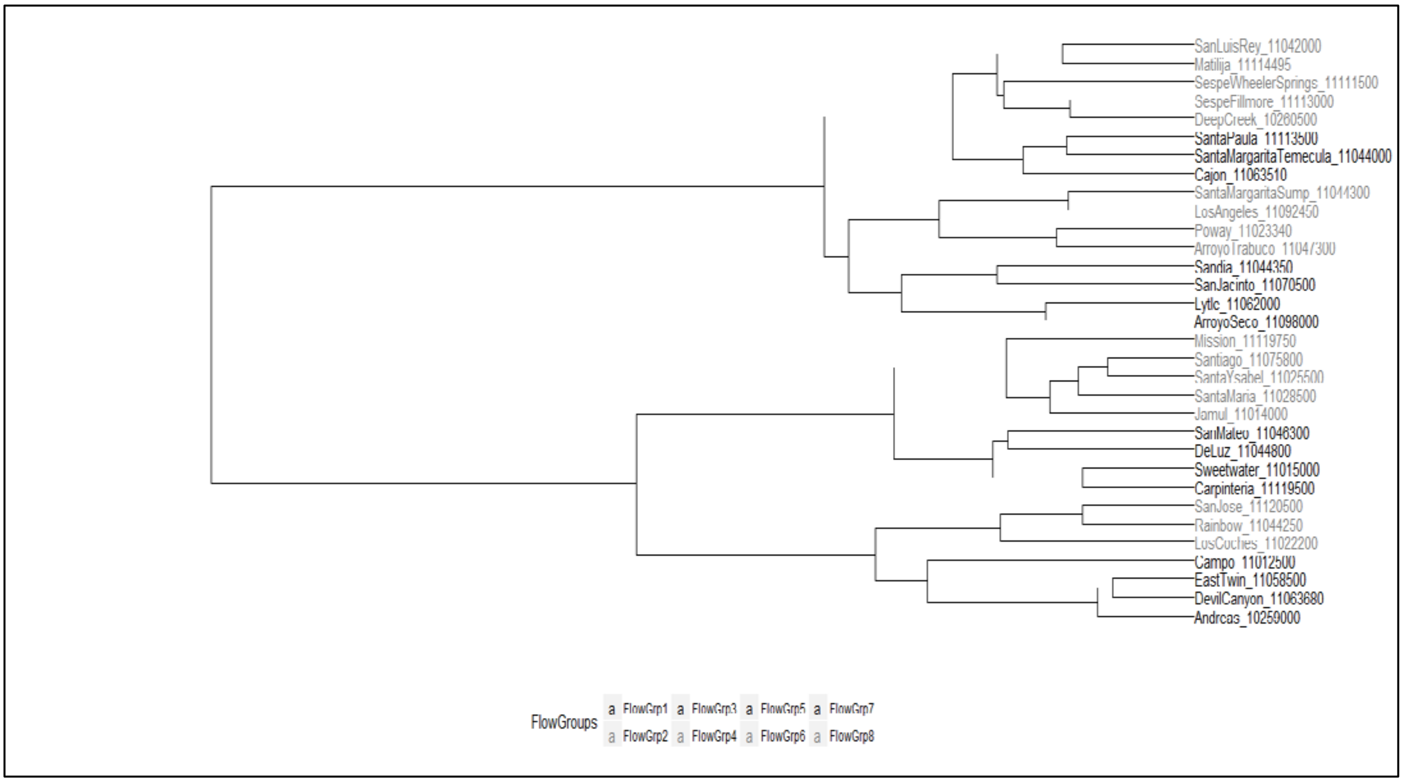

Figure 3-4. Cluster analysis showing 8 groups of hydrologically similar calibration gauges from the ensemble. Subsequently, a random forest model was developed to predict cluster membership of novel sites based on watershed characteristics. This random forest model was used to estimate the statistical proximity between an ungauged site and each calibration gauge. The most proximal gauge was then assigned to an ungauged site.

Gauged models from the ensemble were best matched with between 1 and 64 ungauged sites, with the top five models in the ensemble being matched to 46% of the 572 ungauged sites (Table 3-5). These five gauged models are representative of a wide range of watershed area and imperviousness. Campo, Lytle and Devil Canyon, which were poorly performing during calibration and validation, were assigned a total of 26 ungauged sites, which comprised only 4% of the total ungauged sites (n= 572).

Table 3-5. Number of ungauged sites assigned to each gauged model

Site Name	NSE	Percent Error	Percent Error RBI
Andreas	0.58	5.3	8.1
Arroyo Seco	0.42	19.4	5.2
Arroyo Trabuco	0.73	15.7	33.2
Campo	0.49	7.4	55.3
Carpinteria	0.83	0.3	0.1
DeLuz	0.9	6.4	8.3
Sweetwater	0.57	2.9	71
Devil Canyon	0.57	59.5	27.8
East Twin	0.4	33.4	4.9
Jamul	0.46	15.8	13.8
Los Angeles	0.95	0	23.1
Los Coches	0.79	4.5	22.2
Poway	0.91	18.8	20.7
Lytle	0.44	18.9	6.3
Matilija	0.84	16.7	18.8
Mission	0.83	8	26.6
Rainbow	0.73	10.7	8.3
San Jose	0.81	5.5	25
San Mateo	0.75	2.5	50.7
Santa Maria	0.73	1.3	23.4
Santa Paula	0.53	0	17.8
Santa Ysabel	0.72	3.4	16
Santiago	0.58	0.1	27.6
Sespe Fillmore	0.61	0	22.6
Sespe Wheeler Springs	0.58	5.2	47.2
Sandia	0.63	0	24

3.3.4 Ability to predict flows at ungauged sites

Flow predictions at ungauged sites generally validated well, with the average r2 flow prediction at the 15 validation around 0.45, and for 57% of the validation combination, the r2 values are greater than 0.40. The performance of the HEC-HMS models was primarily dictated by three factors, the quality of the precipitation estimates, model parameters assigned, and presence of flow control or diversion structures. The impact of precipitation on model performance is observed in the interannual variation at a given site (Table 3-6), i.e., for the site 1, the worst year (1996) has a 0 fit, but the best year (1998) has a 0.96 fit. There is some bias in the prediction, with the predicted values tending to be higher than the observed for sites where bias is observed (Figure 3-5, for site 14, best prediction r2 = 0.63, and slope = 13.92).

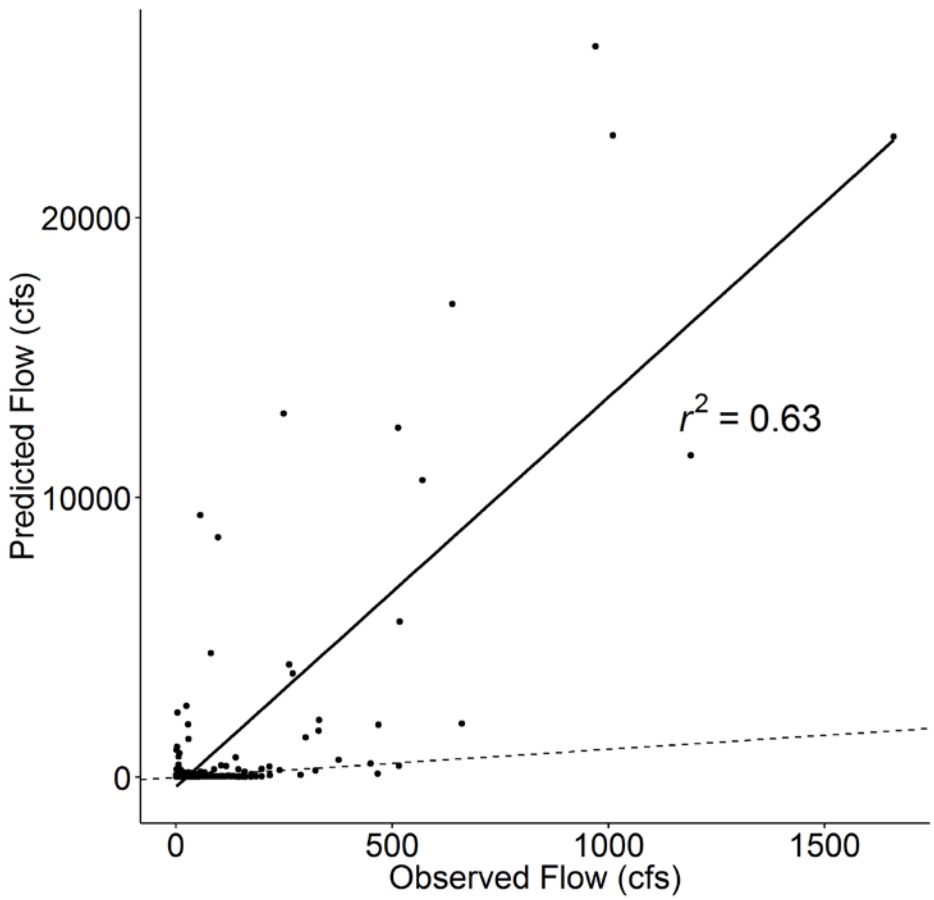

Figure 3-5. Predicted versus observed flows for site 14, matched to Los Angeles model parameter. Predictions are for the year 1993, where the dashed line is 1:1 ratio.

Table 3-6. R-squared (annual) values for validation gauges, the table shows worst, best, and average r2 values, along with p-value and slope. Values highlighted in green have good performance.

Validation Sites	Worst Year Rsq	Slope	Best Year Rsq	Slope	Average Rsq
Site 1	0	0.21	0.96	1.62	0.21
Site 2	0	0	0.08	0.07	0.03
Site 3	0.14	0.49	0.50	0.42	0.32
Site 4	0.29	-0.12	0.73	1.6	0.49
Site 5	0.37	0.24	0.95	0.61	0.72
Site 6	0.03	0.08	0.52	0	0.14
Site 7	0.02	0.16	0.52	0.49	0.24
Site 8	0.01	0.02	0.81	0.78	0.52
Site 9	0.03	0.09	0.44	0.51	0.21
Site 10	0.12	0.19	0.16	0.22	0.14
Site 11	0.49	0.7	0.71	0.25	0.61
Site 12	0.54	1.14	0.91	1.62	0.67
Site 13	0.64	0.47	0.86	0.29	0.77
Site 14	0.08	1.98	0.63	13.92	0.31
Site 15	0.27	0.01	0.95	0.01	0.62

3.3.5 Ability to predict metrics

In general, the metrics calibrated and validated well, but the validation results varied by metric category (Table 3-7) for the four precipitation conditions (overall, dry, wet, and average). Metric-precipitation combinations that validated poorly (i.e., $r^2 < 0.25$) at the calibration sites were excluded from further analysis. Magnitude metrics tended to calibrate and validate the best, with r^2 values are high as 0.99. There was a drop in the performance between the calibration gauges and the validation gauges for timing metrics. For the 7 timing metrics, there were 23 metric:precipitation combinations out of 27 possible combinations with r^2values higher than 0.25. However, for the validation gauges, only 4 combinations out of a possible 28 had an r^2 value higher than 0.25. For all metric:precipitation combinations, the models tended to predict dry year metrics better than the overall, wet or average years.

Table 3-7. Metric validation by category at the calibration gauges (N=26) and validation gauges (N= 15). There are five main categories of metrics, under each category metric in were removed from further analysis due to poor performance. The metrics with asterisks have dual flow threshold issue. Values of $r^2 < 0.25$ are highlighted.

Metrics			Calibration Gauges				Validation Gauges			
			Overall	Dry	Average	Wet	Overall	Dry	Average	Wet
Duration										
	LowDur *	days/event	0.54	0.51	0.32	0.22	0.33	0.43	0.4	0.37
	HighDur *	days/event	0.25	0.26	0.1	0.47	0.11	0.09	0.09	0.4
	NoDisturb	days/year	0.33	0.44	0.38	0.37	0.33	0.43	0.43	0.36
	Hydroperiod	proportion	0.49	0.48	0.38	0.2	0.33	0.47	0.3	0.61
	Per_Low Flow	proportion	0.96	0.92	0.89	0.49	0.01	0.1	0.11	0.1
Frequency										
	HighNum *	events/year	0.7	0.33	0.47	0.7	0.51	0.34	0.31	0.37
	FracYearsNoFlow	proportion	0.33	0.25	0.1	0.28	0.08	0.2	0.05	0
	LowNum *		-0.04	0.21	-0.04	0.01	0.004	0.16	0.001	0.05
	MedianNoFlowDay	days/year	0.37	0.57	0.37	0.2	0.34	0.5	0.31	0
Magnitude										
	MaxMonthQ	cms	0.69	0.95	0.83	0.82	0.6	0.11	0.14	0.69
	MinMonthQ	cms	0.39	0.34	0.24	0.03	0	0	0	0
	Q01	cms	0.99	0.99	0.75	0.99	0.07	0.14	0.05	0.1
	Q05	cms	0.99	0.99	0.73	0.89	0.16	0.32	0.05	0.08
	Q10	cms	0.99	0.99	0.71	0.72	0.25	0.5	0.04	0.07
	Q25	cms	0.97	1	0.73	0.31	0.4	0.73	0.02	0.03
	Q50	cms	0.69	0.97	0.68	0.27	0.37	0.8	0.02	0.01
	Q75	cms	0.56	0.97	0.74	0.7	0.21	0.82	0.12	0.1
	Q90	cms	0.71	0.92	0.72	0.73	0.55	0.81	0.4	0.55
	Q95	cms	0.77	0.9	0.45	0.8	0.66	0.71	0.63	0.65
	Q99	cms	0.94	0.98	0.95	0.86	0.23	0.51	0.02	0.15
	Qmax	cms	0.91	0.95	0.91	0.85	0.4	0.67	0	0.34
	Qmean	cms	0.9	0.99	0.9	0.89	0.4	0.74	0.31	0.61
	QmeanMEDIAN	cms	0.9	0.99	0.9	0.89	0.43	0.74	0.61	0.31
	Qmed	cms	0.57	0.99	0.68	0.27	0.59	0.76	0.37	0
	Qmin	cms	0.99	1	0.74	0.44	0.83	0.92	0.08	0.07
Timing										
	C_C	ratio	0.81	0.66	0.57	0.33	0.09	0.17	0.09	0.05
	C_CP	ratio	0.81	0.63	0.55	0.41	0.01	0.02	0.04	0.07
	C_M	ratio	0.76	0.61	0.54	0.46	0.04	0.01	0.01	0.02
	C_MP	ratio	0.81	0.62	0.54	0.41	0.01	0	0.04	0.07
	C_P	ratio	0.71	0.61	0.41	0.33	0.28	0.28	0.39	0.31
	MinMonth	month	0.24	0.34	0.39	0.03	0.25	0.08	0.08	0.16
	MaxMonth	month	0	-0.04	0.19	0.77	0.08	0.01	0.25	0.01
Variability										
	RBI	unitless	0.72	0.64	0.32	0.72	0.06	0.03	0.02	0.07
	SFR	proportion	-0.01	0.19	0.75	-0.03	0.8	0.05	0.84	0.51
	PDC50		-0.04	0.14	0.25	-0.04	0.06	0.06	0.07	0.006
	BFR		-0.04	0.05	0.18	-0.04	0.02	0.006	0.04	0.75
	QminIDR	cms	0.71	NA	NA	NA	0	0.03	0	0
	QmeanIDR	cms	0.88	NA	NA	NA	0.45	0.28	0.95	0.37
	QmaxIDR	cms	0.85	NA	NA	NA	0.65	0.53	0.09	0.64

3.3.6 Regional Alteration

The influence of anthropogenic actions on flow alteration varies by metric category (Table 3-8). Magnitude metrics tend to increase in response to urban and agricultural land uses, whereas the timing and duration metrics are mostly unchanged. We observed a decrease in the duration metrics under agricultural and urban land use, especially the number of no disturbance days and percent low flow days. Duration metrics are mostly unchanged in the streams in undeveloped areas. Agricultural and urban land use causes a decrease in the timing metrics. Alteration in the streams in the three main stream classes (1, 2 and 4) are mixed. For stream class 4, which comprises of Southern California's large, lowland rivers, most magnitude metrics increase, especially the high flows, and most stream kms are relatively unchanged for the duration and timing metrics. For stream class 1, representative of the high elevation mountain streams the receives snowmelt, majority of the stream kms show an either an increase or no change in the magnitude metrics. Similarly, for the timing and duration metrics, the stream kms in class 1 are relatively unchanged. Finally, class 2, which represents lower elevation stream driven mostly by rainfall and groundwater, most stream kms show an increase in the magnitude metrics, and no change for the timing and duration metric categories.

Table 3-8. Increasing and decreasing trends in 5 categories of regional metrics, by land use (agriculture, undeveloped and urban), and by major stream class in the region (classes 1,2 and 4).

	Region			Land Use									Hydrologic Class								
				Agricultural			Undeveloped			Urban			Class 1			Class 2			Class 4		
Metric	Dec	NC	Inc	Dec	NC	Inc	Dec	NC	Inc	Dec	NC	Inc	Dec	NC	Inc	Dec	NC	Inc	Dec	NC	Inc
Duration																					
HighDur	30	43	27	45	20	36	10	57	33	72	5	22	15	49	36	33	52	15	33	31	36
Hydroperiod	1	71	27	5	52	44	1	78	22	1	46	53	0	83	16	2	68	29	0	67	33
LowDur	18	62	19	18	45	37	10	78	11	31	20	49	21	73	7	20	56	24	17	65	18
NoDisturb	55	35	10	89	8	3	37	44	18	91	5	3	44	32	24	56	39	5	64	27	8
Per_LowFlow	42	47	10	86	10	4	34	62	3	68	18	15	30	59	11	44	48	8	53	43	4
Frequency																					
HighNum	0	56	44	2	6	92	0	79	21	0	4	96	0	69	31	0	61	39	0	39	61
MedianNoFlowDays	22	78	0	22	78	0	15	85	0	38	62	0	12	88	0	27	73	0	21	79	0
Magnitude																					
MaxMonthQ	5	3	91	1	3	96	15	7	78	0	0	100	13	3	84	4	2	95	1	6	93
MinMonthQ	36	23	41	27	6	67	44	27	29	34	12	54	40	33	27	41	10	49	18	37	46
Q01	32	57	11	36	31	33	32	52	16	51	44	6	18	65	17	38	54	8	39	55	6
Q05	20	47	34	16	38	46	17	55	28	40	47	13	9	66	25	20	33	47	17	53	30
Q10	31	47	22	35	53	12	32	39	28	51	38	10	16	48	35	38	47	14	34	53	13
Q25	38	32	30	34	21	45	35	33	33	51	23	26	19	45	37	58	21	22	24	35	42
Q50	29	22	49	12	5	83	25	24	51	37	6	56	16	46	38	43	7	50	14	30	56
Q75	47	17	36	32	6	62	40	27	33	43	4	53	27	29	44	66	7	27	26	29	45
Q90	38	11	51	33	5	63	39	19	42	22	4	74	34	16	50	53	5	42	20	26	55
Q95	32	15	54	22	3	76	39	25	36	7	2	92	20	39	41	46	6	49	20	12	68
Q99	20	20	60	8	3	89	40	26	34	1	6	94	24	40	36	26	6	69	14	23	63
Qmax	3	2	95	10	3	88	8	5	86	0	0	99	7	2	91	2	2	97	2	5	94
Qmean	2	0	98	0	3	97	7	1	92	0	0	100	4	0	95	2	0	98	0	1	99
QmeanMEDIAN	59	2	39	35	3	63	74	3	24	11	4	85	75	2	22	55	1	44	58	3	39
Qmed	49	2	49	56	4	40	36	4	60	63	6	32	24	2	73	66	2	32	38	4	58
Qmin	58	12	29	44	12	44	35	11	54	74	19	7	44	12	44	66	11	22	49	15	36
Timing																					
C_C	35	58	7	63	29	8	17	71	12	92	7	2	18	75	7	36	59	5	40	54	6
C_CP	42	52	7	65	22	13	20	63	17	94	4	2	25	66	9	46	51	3	43	48	8
C_M	8	53	38	41	20	40	16	68	16	5	6	89	10	68	22	3	54	43	10	48	41
C_MP	7	51	42	15	20	65	18	62	20	2	5	93	10	65	25	3	51	45	8	48	45
C_P	21	67	12	40	45	15	7	90	3	59	16	25	7	86	7	24	65	11	20	63	16
MaxMonth	7	86	7	8	78	15	0	92	8	22	68	10	1	93	5	5	91	4	9	85	6
MinMonth	12	79	8	5	90	5	2	97	2	30	45	26	3	94	3	15	75	10	16	76	8
Variability																					
QmaxIDR	9	5	86	24	3	73	13	6	81	13	0	86	19	11	70	6	2	92	7	6	87
QmeanIDR	26	0	74	5	3	93	19	1	80	3	0	97	27	1	71	31	0	69	22	0	77
QminIDR	67	26	6	62	34	4	52	31	17	78	17	5	55	40	6	77	15	8	60	34	6
RBI	19	3	78	4	3	93	28	4	68	0	4	96	40	4	56	7	1	92	29	4	67
SFR	75	20	6	94	6	0	57	21	22	93	6	1	54	36	10	90	6	4	67	27	6

3.4 Discussion

Approximately 79% of the region shows some degree of hydrologic alteration, and approximately 40% of the sites can be considered severely altered with at least 10 metrics in the top fourth quartile of hydrologic alteration (Figure 3-6). Among the five metric categories (timing, frequency, magnitude, duration and variability), the magnitude metrics are usually the most altered at the severely altered sites (number of altered metrics in the top 25^{th} quartile > 10). This is comparable to Carlisle et al (2010) which found that 86% of the assessed streams in conterminous United States were altered for magnitude metrics. Contrary to Carlisle et al. (2010) we see an inflation or overall increase in the high flow metrics (Q99) under wet and average conditions. This effect could be regional and connected to higher imperviousness in the catchments.

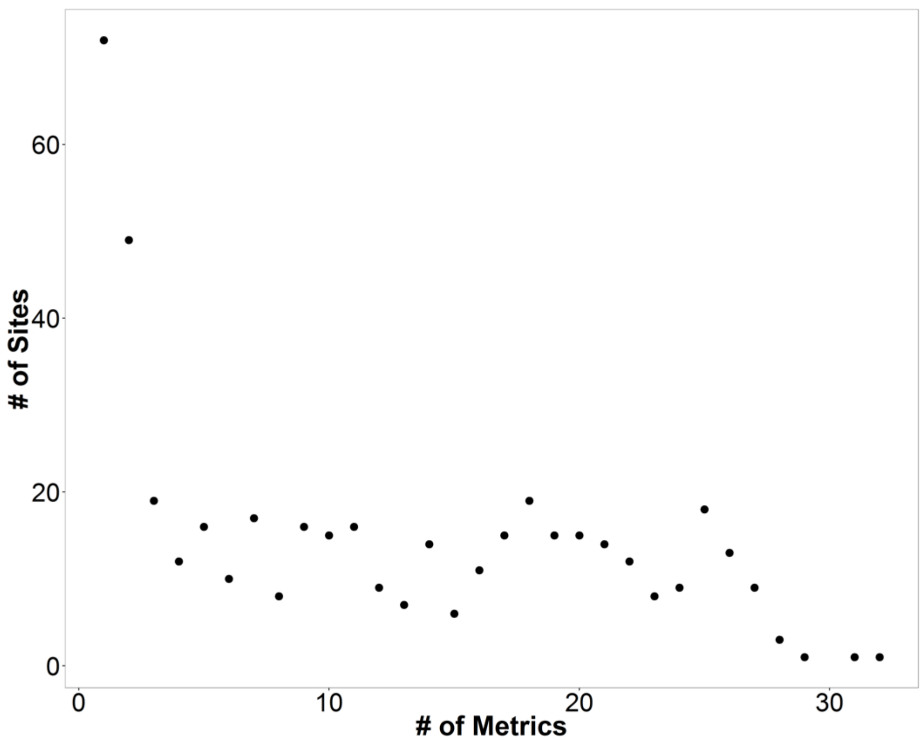

Figure 3-6. Number of severely altered metrics by number of sites with delta H values in the top quartile. At 572 ungauged sites, delta H for the 36 metrics was estimated and grouped into quartiles. At each site the number of metrics that show alteration in the top quartile are counted and presented on the x-axis. The count of number of sites are presented on the y-axis.

The degree of hydrologic alteration varies between the wet, dry and average years, with higher degree of alteration in the average or the dry years compared to wetter years for 35 of the 36 metrics. This variation can have management implications, for example, a site could be considered hydrologically altered in the dry years but not in the wet years. This interannual variability may partially mask patterns in overall alteration and suggests that management actions should be tailored based on climatic conditions.

Patterns of increased magnitude metrics and decreased duration and timing metrics associated with anthropogenic land use are consistently observed throughout the region. This effect of increasing streamflow due to agricultural practices has been reported in other studies (Raymond et al 2008), and for other land use (Stohlgren et al. 1998, Yan et al 2013). Further analysis of the effect of increasing imperviousness on the metrics shows the imperviousness is positively correlated with increasing alteration. Comparing the degree of alteration for three representative metrics, LowDur (duration), Qmean (magnitude) and QmaxIDR (variability) shows that hydrologic alteration is pervasive in catchments (48% of sites) with impervious cover higher than 5% (Figure 3-7). Hydrologic responses and biological responses at such low levels of imperviousness is unusual compared to thresholds of around 8-10% observed in other studies (Wang) with changing imperviousness at the 584 sites.

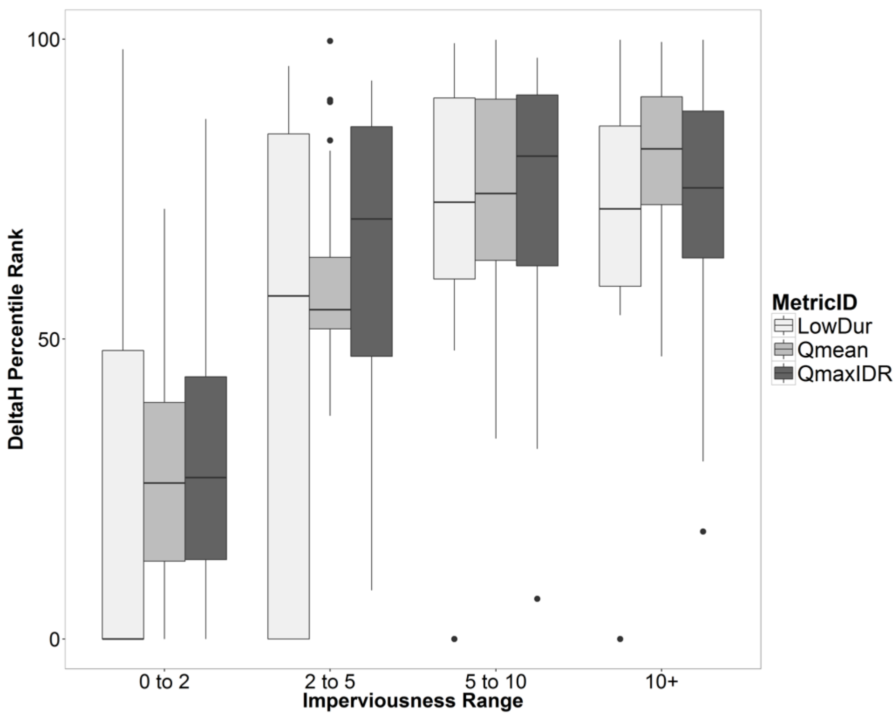

Figure 3-7. Delta H for three selected metrics LowDur (duration), Qmean (magnitude), and QmaxIDR (variability) with changing imperviousness at the 572 sites. Horizontal box lines, from lower to upper, represent 25th, 50th and 75th percentiles. Whiskers lengths are 1.5 multiplied by the Interquartile range.

3.4.1 Advantage of mechanistic approach for estimating flow metrics at ungauged sites

Application of a mechanistic model ensemble to predict regional hydrologic alteration provides some advantages over statistical methods that are typically applied at the regional scale. Because this approach is based on physical processes it allows for consideration of a broad suite of flow metrics that are derived from hourly flow data and represent all aspects of the hydrograph. For example, our model ensemble produced metrics that validated well for four of the broad flow categories: duration, frequency, magnitude, and variability. Statistical methods such as regression based models (Carlisle 2010), neural networks (Besaw et al. 2010), and flow duration curves (Holmes et al. 2002), provide a static flow characterization at ungauged sites for a pre-determined set of flow metrics (i.e. whatever is modeled). The mechanistic approach generates continuous granular flow data (hourly time step) at the ungauged locations allowing for consideration of metrics that may be applicable to a variety of ecological endpoints (e.g. fish vs. invertebrates), different life history requirements (e.g. breeding vs. migrations), and consideration of a range of management tradeoffs (e.g. diversions or discharges). Moreover, once the models are established, they can be applied to new management questions or locations in a straightforward manner. The sub-daily flow data can be useful in managing flow regimes to maintain ecological function (Richter et al., 1997, 2003; Poff et al., 2003), especially in arid regions such as Southern California, where precipitation patterns are extremely variable with short rainstorm events lasting a couple hours, typically during the winter season (Gasith and Resh 1999, Nezlin and Stein 2005). The regional ensemble allows for rapid application to new sites of interest with minimal effort on model parameterization and no additional calibration or validation requirements. This approach can allow managers to explore the impact of land use conversion in a part of the catchment on the receiving waters. Similarly, these models can be easily adapted to site best management practices to manage alteration on an ongoing basis

3.4.2 Lessons learned

The three calibration criteria selected in this study emphasize different components of the flow regime, especially low flow frequency and flashiness. Relying on a simple overall fit, as depicted by NSE is insufficient for representing low flow periods, intermittency and flashiness in these Southern California streams, for example, the models calibrated for only NSE tend to have high error for low flow days. However, multi-objective calibration comes with its set of caveats (Price et al. 2012), and can result in a decrease in the NSE values. For example, at the Lytle Creek site, the NSE value decreases from 0.78 (single objective calibration for NSE) to 0.42 in the multi-objective calibration model. However, the lower overall NSE is likely reflective of our level of confidence in model performance of a range of ecologically relevant flow conditions.

The performance of the HEC-HMS models was primarily dictated by the quality of precipitation estimates, precipitation patterns, model parameters assigned, and presence of flow control or diversion structures. As expected, the model predictions were better for years with good precipitation data inputs compared to years with missing data. The predictions were also affected at sites with extreme topography and large orographic effects. Challenges of predicting flow in relatively steep streams with fast rising hydrographs can be addressed by selecting a different routing method, such as the Muskingum-Cunge in HEC-HMS. However, we chose not to do so since there is a possibility that these gauged models will get assigned to an ungauged site which is perhaps located in milder topography. In the future, this can be addressed by weighting the slope higher during the clustering and model assignment process. Finally, watersheds with flow diversion or other hydrologic control require additional adjustments, such as changing baseflow or losses to the assigned HEC-HMS models.

The mechanistic model ensemble provides a calibrated template that represents hydrological processes in different catchments within a region. Therefore, we based the model assignment process on similarities in hydrologic properties between the gauged and the ungauged sites. We believe this approach is reasonable and parsimonious for regional application. However, alternative approaches are possible (or have been used by others) and should be evaluated in future studies. Our approach represents a deviation from other model assignments which rely on the similarities in the physical properties of the catchments. An alternate approach of assigning a gauged model from the ensemble to an ungauged site could be by clustering the gauged models based on the errors from the jack-knife validation exercise, where the error matrix is used as a dissimilarity matrix in cluster analysis rather than clustering them based on observed flow metrics. This approach may be better at identifying mutually transferable pairs of models, rather than hydrologically similar models. Additionally, the model assignment was largely based on stressor gradients, even though variables related to natural factors had a chance for selection. Given that our goal was to use these models to simulate both current and historic conditions, it may be reasonable to restrict model selection to natural factors. Finally, we selected just a single HEC-HMS model, rather than a combination of several models, perhaps weighting by proximity. These alternative approaches to model extrapolation should be explored through future studies to determine if adding complexity provide measurable benefit in model performance at assigned ungauged sites.

This study discovered and addressed difficulties with using duration metrics which may involve a shifting baseline, e.g. high duration, low duration, number of no disturbance flows, and number of high flow events. These metrics are based on benchmark discharge values used to identify high- or low-flow events (typically, the 90^{th} or 10^{th} percentile). Comparisons between historic and current conditions can be complicated and counter-intuitive when this benchmark changes dramatically (e.g. under certain conditions, large reductions of flow may appear to increase the frequency and duration of high-flow events). An example is illustrated in Figure 3-8, where the Highdur values estimated for historic conditions is 16 days. When the Highdur values are estimated for current conditions using the new Q90 (current) threshold the number of days reduces to 15. This indicates that the high flow days are lower under the current conditions even though the catchment has undergone landuse change; a finding that seems counter-intuitive. When Highdur is estimated for current conditions using the Q90 historic threshold, the number of days with high flow increases to approximately 38. Therefore, management planning based on metrics estimated using moving thresholds can be misleading. This issue may not have been apparent in past use of similar flow metrics because those studies were not attempting to simulate historic conditions or were relying on more punctuated alterations, such as construction of a dam, where this difference may be less apparent. To the best of our knowledge no other study has reported issues with shifting benchmarks, and we propose that the alternative approach applied in this study of using single historic benchmark for both current and historic condition metric estimation balances out the ‘changing baseline’ issue.

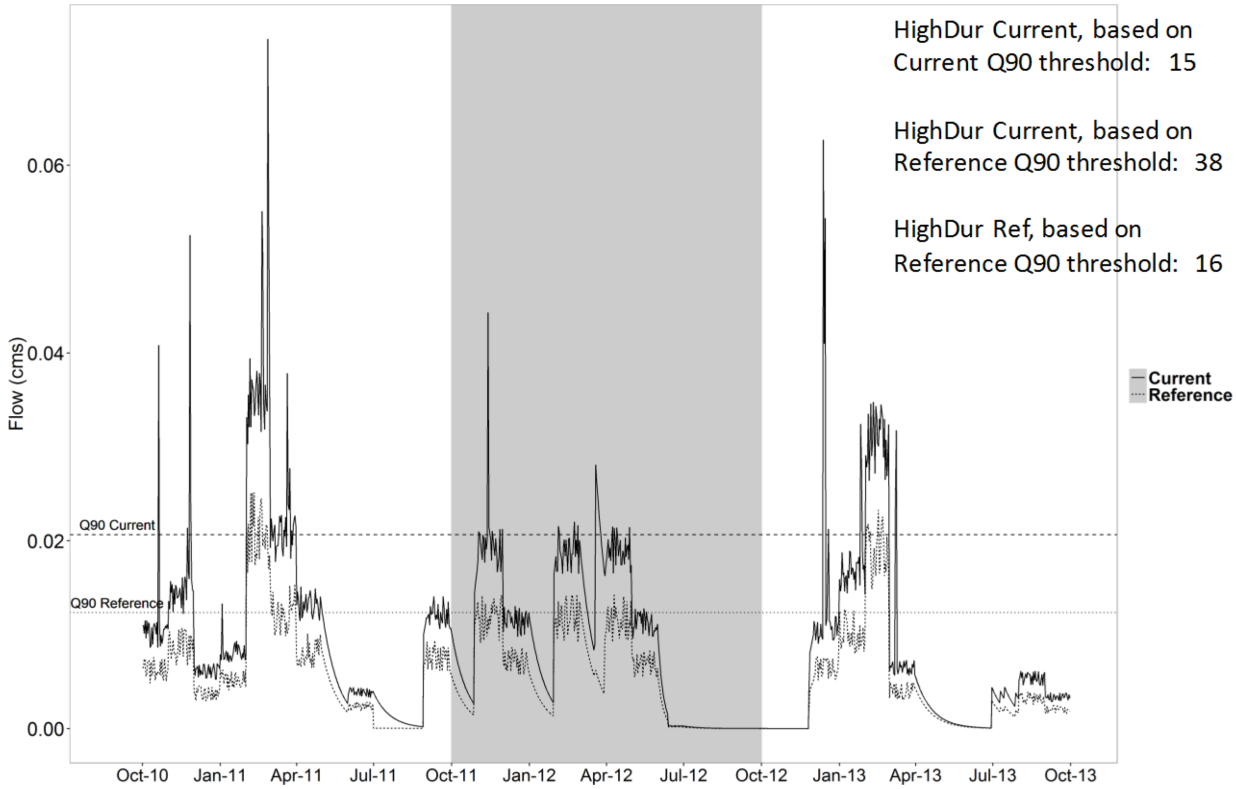

Figure 3-8. Comparison of current and reference flow for a sample bioassessment site showing the effect of use of different thresholds. Conclusions about changes in duration of high flow events would vary dramatically if only a single threshold based on reference is issued vs. different thresholds were used for current and reference conditions.

3.4.3 Conclusion and future work

Our goal was to provide a regional understanding of the hydrologic alteration, and to combine this understanding with multiple robust biological datasets to develop a broad suite of flow ecology relationships to support management decisions. The modeling approach we described successfully predicted the flows, and a range of flow metrics at many sites spread over a wide geographical region. Relative ease of transferability and applicability makes this a useful tool in the region for new sites and scenarios. A distinct advantage of the mechanistic approach is the ability to generate site specific scenarios, such as response to implementation of stormwater capture structures, or rapid urban development in a given catchment. This will aid in understanding the implications of the regional flow ecology relationships for specific management applications.

In the future, we anticipate the application of flow-ecology relationships in predicting changes in the hydrologic regimes under various management options and climate change, and developing scenarios and risk analysis (Poff et al., 2003; Stewardson & Gippel, 2003; Richter et al., 2006).

This is particularly applicable to Southern California given the impetus to reuse and recycle treated effluent and stormwater in the region (California Water Action Plan 2015, Hering et al 2013). Impacts of climate change in Southern California will manifest in form of flooding, and shifts in precipitation patterns (Hanak and Lund 2012). Foreseeing the impact of these factors in the hydrologic regimes in the region and proactively developing management strategies to mitigate the impacts will inform future decisions regarding complex water management issues.

4.0 EVALUATION OF BIOLOGICAL CONDITION RELATIVE TO FLOW TARGETS

4.1 Background

Indicators of biological integrity, such as benthic macroinvertebrates, can serve as an ultimate measure of the impacts of hydrologic alteration because these assemblages integrate the totality of stresses to which they are exposed over time (Karr and Chu 2000). Indices of stream condition based on benthic macroinvertebrates are widespread in both monitoring and regulatory programs, and are increasingly used to set management objectives (e.g., US EPA 1990). Understanding how flow alteration affects biological indices enables managers to establish flow management targets, identify when flow is a predominant stressor affecting biological condition, plan restorations that help recover biological condition, or avoid activities that lead to degradation.

Establishing flow ecology targets based on benthic invertebrate indices requires large data sets of both biological and hydrological condition that can be used to derive relationships that are applicable to streams across broad ranges of conditions within a given stream class or region. In Southern California, we were able to build off approximately eight years of regional biomonitoring (supported through both state and regional programs) that have generated approximately 600 probabilistically sampled bioassessment sites. The regional model ensemble described in the previous chapter provides the ability to estimate hydrologic alteration at most of these sites, providing a large data set from which we can develop flow ecology relationships.

Our objectives were to evaluate responses in indicators of biological health (specifically the California Stream Condition Index [CSCI] and its components, Mazor et al. 2016) to measures of hydrologic alteration using logistic regression. We then used these relationships to set flow targets that could be applied to ungauged sites throughout Southern California. We then created an index to rank metrics based on the strength of their association with biological condition determined by boosted regression trees, selecting metrics that represent different components of the hydrograph. This index was then applied to a probabilistically sampled data set to estimate the linear extent of hydrologically altered streams in Southern California. In conjunction with biological data, the index was used to prioritize management actions and perform rapid causal assessments at a regional scale. Finally, we evaluated the interactive effects of hydrologic alteration, water quality, and habitat degradation through graphical methods and ordination.

4.2 Methods

4.2.1 Estimation of biological alteration

Bioassessment data were collected at 572 unique sites in Southern California under a variety of programs, (most under regional stream survey of the Stormwater Monitoring Coalition [SMC], Mazor 2015) (Figure 4-1). Benthic macroinvertebrates were sampled according to Ode (2007), and scored with the California Stream Condition Index (CSCI) following Mazor et al. (2016). The CSCI is a predictive index that compares observed taxa and metrics to values expected under reference conditions based on site-specific landscape-scale environmental variables, such as watershed area, geology, and climate. It includes two components: a ratio of observed-to-expected taxa (O/E), and a predictive multi-metric index (MMI) made up of 6 metrics related to ecological structure and function of the benthic macroinvertebrate assemblage. Because the CSCI and all of its components are based on site-specific reference expectations, they are minimally influenced by major natural gradients, and can therefore be used as a measure of biological alteration under anthropogenic stress. CSCI scores and all components were classified as indicating "intact" or "altered" condition, using the normal approximation of the 10th percentile of CSCI reference calibration scores as a threshold. For the CSCI, O/E, and MMI, these thresholds are published in Mazor et al. (2016). For the 6 biological metrics, thresholds were also calculated as normal approximation of the 10th percentile of reference calibration values, based on the means and standard deviations reported in Mazor et al. (2016); all biological thresholds are presented in Table 4-1.

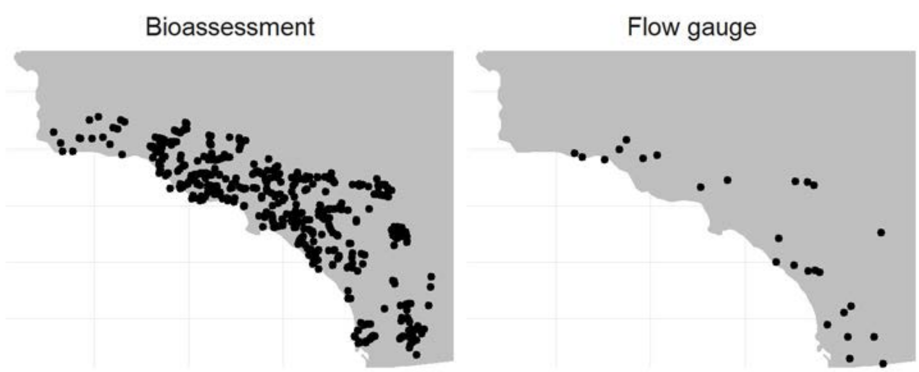

Figure 4-1. Locations of bioassessment sites and flow gauges used to develop models. Inset shows the study area within California.

Table 4-1. Thresholds based on the normal approximation of the 10th percentile of reference calibration scores used to develop the California Stream Condition Index (CSCI, Mazor et al. 2016). Approximations were calculated from the reference mean and standard deviation using the qnorm function in R (R Core Team, 2016). MMI: Multi-metric index. O/E: Observed-to-Expected taxa. EPT: Ephemeroptera, Plecoptera, and Trichoptera.

Biological response variable	Mean	Standard	Threshold	Percent above
CSCI	1	0.16	0.79	43
MMI	1	0.18	0.77	34
O/E	1	0.19	0.76	58
Clinger Percent Taxa score	0.72	0.17	0.49	36
Coleoptera Percent Taxa score	0.60	0.21	0.32	42
EPT Percent Taxa score	0.74	0.16	0.54	29
Intolerant Percent score	0.47	0.25	0.15	97
Shredder Taxa score	0.54	0.24	0.23	90
Taxonomic Richness score	0.67	0.20	0.41	72

4.2.2 Estimation of hydrologic alteration

An ensemble of hydrologic models developed at 26 calibration gauges (Figure 4-1) was used to estimate hourly hydrographs under current and historic conditions following Sengupta et al. (in review and previous chapter). Estimates were based on 6 years of rainfall: 2 wet, 2 average, and 2 dry years. Sites where suitable precipitation data could not be estimated were excluded from analysis. Hourly hydrographs were then aggregated to daily discharge, and a suite of flow metrics (Table 4-2) with presumed biological relevance were calculated for both current and reference conditions. Metrics were calculated with all 6 years to estimate the metric under overall precipitation conditions, and also for the 2 years corresponding to each precipitation condition. Metric-precipitation condition combinations that validated poorly (i.e., $r^2<0.25$) were dropped from analysis, yielding a total of 37 metrics and 121 metric-precipitation combinations for analysis. For each metric-precipitation combination, hydrologic alteration was characterized as differences between current and reference condition (see previous chapter).

Table 4-2. Flow metrics, descriptions, and results of analysis. Precipitation conditions are indicated by columns marked O (overall), W (wet), A (average), and D (dry). Gray cells indicate that the metric was analyzed under the indicated precipitation conditions. Solid black dots indicate that the indicated precipitation condition had greater importance for predicting biological endpoints than other precipitation conditions of the given metric. Hollow dots indicate that the metric-precipitation condition combination was selected for inclusion in the index of hydrologic alteration. Imp: Average importance of the metric-precipitation condition combination indicated with a solid or hollow dot. in predicting biological endpoints. Dec: Decreasing target. Inc: Increasing target. ND: No data (i.e., <30 sites along stressor gradient). NS: No significant relationship in the expected direction.

Metric	Unit	Description	O	W	A	D		Imp	Dec	Inc
Duration										
HighDur	days/event	Median annual longest number of consecutive days that flow was greater than the high flow threshold		○				32	-2.9	24
Hydroperiod	proportion	Fraction of period of analysis with flows	•					79	ND	0
LowDur	days/event	Median annual longest number of consecutive days that flow was less than or equal to the low flow threshold			•			48	-69	2.3
NoDisturb	days	Median annual longest number of consecutive days that flow between the low and high flow threshold			○			22	-64	NS
Per_Low Flow	proportion	Percent of time with flow below 0.0283 cms	•					68	-2.7	0.3
Frequency										
FracYearsNoFlow	proportion	Fraction of years with at least one no-flow day	•					118	ND	ND
HighNum	events/year	Median annual number of continuous events that flow was greater than the high flow threshold				○		32	ND	2.9
Median NoFlowDays	days/year	Median annual number of no-flow days			•			91	-212	ND
Magnitude										
MaxMonthQ	cms	Maximum mean monthly streamflow		○				4	NS	1.3

Metric	Unit	Description	O	W	A	D		Imp	Dec	Inc
MinMonthQ	cms	Minimum mean monthly streamflow	•					36	-0	0.2
Q01	cms	1st percentile of daily streamflow			•			64	-0	NS
Q05	cms	5th percentile of daily streamflow		•				53	-0	0
Q10	cms	10th percentile of daily streamflow			•			54	-0	NS
Q25	cms	25th percentile of daily streamflow			•			45	-0	NS
Q50	cms	50th percentile of daily streamflow			•			45	-0	0.4
Q75	cms	75th percentile of daily streamflow		•				35	-0	0.5
Q90	cms	90th percentile of daily streamflow		•				30	-0	4.6
Q95	cms	95th percentile of daily streamflow		•				25	-0	14
Q99	cms	99th percentile of daily streamflow			○			13	-0	32
Qmax	cms	Median annual maximum daily streamflow		•				13	ND	6.3
Qmean	cms	Mean streamflow for the period of analysis		•				13	ND	0.1
QmeanMEDIAN	cms	Median annual mean daily streamflow			•			17	-0.7	1.6
Qmed	cms	Median annual median daily streamflow			•			23	-0.3	NS
Qmin	cms	Median annual minimum daily streamflow			•			31	-0.6	NS
Timing										
C_C	ratio	Colwell's constancy (C) a measure of flow uniformity.				•		88	-0.1	NS
C_CP	ratio	Colwell's maximized constancy (C/P). Likelihood that flow is constant throughout the year	•					76	-0.1	NS
C_M	ratio	Colwell's contingency (M). Repeatability of seasonal patterns.	•					89	-0.1	0

Metric	Unit	Description	O	W	A	D		Imp	Dec	Inc
C_MP	ratio	Colwell's maximized contingency (M/P). Likelihood that the pattern of high and low flow events is repeated across years.	•					62	NS	0.1
C_P	ratio	Colwell's predictability (P=C+M). Likelihood of being able to predict high and low flow events			•			88	-0	0
MaxMonth	month	Month of maximum mean monthly streamflow		•				106	-0.4	NS
MinMonth	month	Month of minimum mean monthly streamflow				•		106	-0.4	1.3
Variability										
QmaxIDR	cms	Difference between 90th and 10th percentiles of annual maxima	○					9	-4.5	2.4
QmeanIDR	Cms	Difference between 90th and 10th percentiles of annual means	•					11	NS	0.1
QminIDR	Cms	Difference between 90th and 10th percentiles of annual minima	•					50	-0	NS
RBI	Unitless	Richard Baker Index (flashiness)				○		10	NS	0.2
SFR	Proportion	90th percentile of percent daily change in streamflow on days when streamflow is rece~~e~~ding (storm-flow recession)			•			18	-0.7	NS

4.2.3 Estimation of water chemistry alteration

Selected analytes commonly sampled along with benthic invertebrate assessments were used as indicators of water chemistry alteration: specific conductance, total Nitrogen, and Chloride. Nutrients and major ions are known to be pervasive contaminants associated with biological degradation in Southern California (Mazor 2015). Specific conductance was measured in the field at 511 sites where bioassessments were conducted. These observed values were compared to values expected under natural conditions based on catchment properties (such as geology and climate) following Olson and Hawkins (2012). Water chemistry alteration was then characterized as log of the ratio of observed to expected specific conductance values. Chloride was available at 243 sites, and total Nitrogen was available at 148 sites.

4.2.4 Calculation of biologically-based targets for hydrologic alteration

We established flow targets for each of the 121 metric-precipitation combinations using logistic regression. The logistic regression produces a probability of specific flow alteration being associated with biological alteration. Thresholds were established separately for hydrologic alteration manifested as in increase in metric value and alteration manifested as a decrease in metric value (Figure 4-2). Gradients that included fewer than 30 hydrologically altered sites were excluded from further analysis. Logistic regressions were used to predict the probability of intact biological condition based on a single measure of hydrologic alteration. The glm function in R, with a binomial error distribution and logit link function, were used for analyses (R Core Team, 2016). Regressions were dropped from further analysis if the coefficient term was not significant ($p>0.05$), or if the relationship was in the wrong direction (i.e., a negative relationship for decreasing gradients, or a positive relationship for increasing gradients).

Logistic regression models were used to predict the likelihood of intact biology being associated with ranges of hydrologic alteration between the minimum predicted metric value to zero change (for decreasing gradients), or from zero to the maximum predicted metric value (for increasing gradients). Relative likelihood of biological response was then calculated by rescaling predictions by the maximum prediction to account for the influence of stressors that degrade biology even when hydrology is unaltered. Targets were then selected as the change in hydrologic metric value (excluding zero) that had the relative likelihood of biological response closest to 0.5, meaning that the likelihood of observing intact biological conditions when the target is exceeded is half the likelihood under hydrologically unaltered conditions. The most conservative targets were then selected for each metric from among the all the biological response variables tested (i.e. CSIC, pMMI, O/E, component metrics). To explore the role of classification, targets derived from the complete regional data set were compared to targets derived for subsets of sites belonging to hydrologic stream classes that were well represented in the region (i.e., Classes 1,2, and 4; $n > 100$).

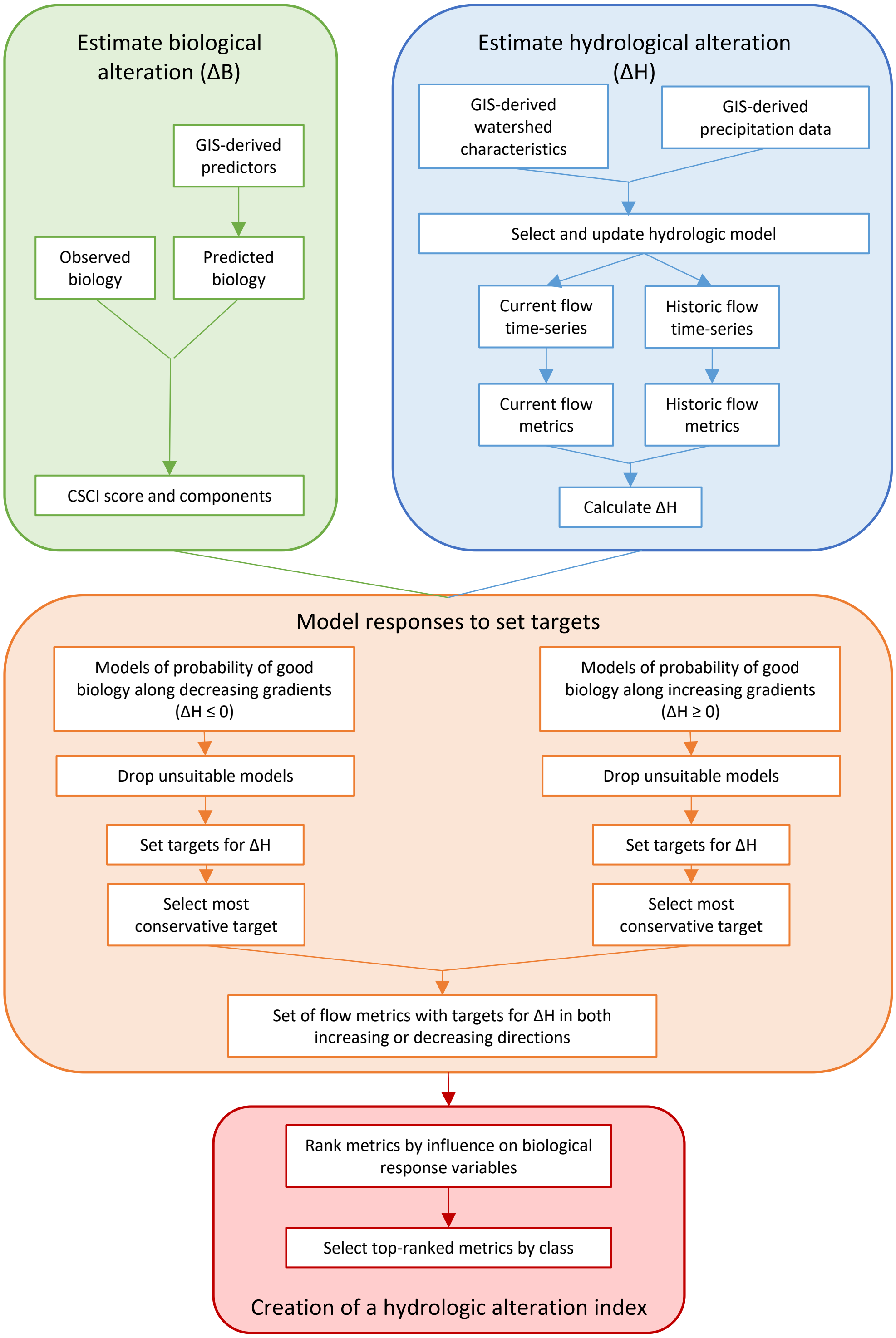

Figure 4-2. Workflow used to calculate, evaluate, and select priority flow metrics for inclusion in the index of hydrologic alteration and for use in establishing regional flow targets. Analyses of biologic alteration are represented in green boxes; analyses of hydrologic alteration are represented in blue boxes

4.2.5 Selection of hydrologic metrics and development of a hydrologic alteration index

We used boosted regression tree (BRT) analysis to rank hydrologic metrics based on their relationships with biological condition for the full suite of 121 flow metric-precipitation condition combinations. BRT models were run using the gbm package in R (Ridgeway 2015) and with specific code from Elith et al. (2008). Each BRT model was developed with the following parameter settings: we used a bag fraction of 0.50, a learning rate of 0.0005 for developing our models, and a tree complexity of 5. Variable relative importance (VRI) was calculated using formulae developed by Friedman (2001) and implemented in the gbm package to estimate the relative influence of each flow metric. Calculations of VRI are based on the number of times a variable is selected for splitting, weighted by the squared improvement to the models as a result of each split, averaged over all trees. VRI values were ranked within in each biotic response model from 1 to 121, with 1 being the best rank. Ranks were then averaged across all 9 biological response variables. Metric-precipitation condition combinations were selected for further analysis if they had at least one target supported by the logistic regression analysis, described above. Within a metric, only the best-ranked precipitation condition was selected for further analysis.

To select a subset of metrics to use in a hydrologic alteration index, up to two metrics were selected in order of average rank from each metric class (i.e. duration, magnitude, variability, frequency), as long as the average rank was better than the median average rank. The subset of flow metrics was re-run in new BRTs in order to evaluate their relationship with biological response variables. Metrics were scored 0 if they met targets, 1 if they failed targets, and 2 if they failed by more than twice the target value. Sites that scored 2 or more were designated as hydrologically altered. To examine the relationship between the index and biological response variables, the index score was then plotted against each response variable. A smoothed fit from general additive models was added by using the default settings of the geom_smooth function in the ggplot2 package in R (Wickham 2009, R Core Team 2016).

4.2.6 Assignment to classes for management priorities

Sites were assigned to one of four classes based on their biological and hydrological condition (Table 4-3). Biological condition was inferred using CSCI scores: Sites with scores greater than 0.79 were designated as biologically intact, and sites with lower scores were designated as biologically altered (Mazor et al. 2016). Hydrologic alteration was inferred using the hydrologic alteration index described above. Hydrologically unaltered (i.e., hydrologic alteration index score = 0) and biologically intact (i.e., CSCI score ≥ 0.79) sites were put into a “protection” class connoting the need to protect these sites from further degradation. Hydrologically altered and biologically altered sites were put into a “monitoring” class containing sites that may be resilient to stressors related to hydrologic alteration and require monitoring to ensure they continue to support biological health. Hydrologically altered and biologically altered sites were put into a “flow management” class; these sites should undergo a causal assessment to determine if flow management is likely to improve biological condition. Hydrologically unaltered and biologically altered sties were put into an “other management class; these sites should also undergo causal assessments with other management options prioritized over flow management.

Table 4-3. Management categories defined based on combination of hydrologic and biologic alteration

	Poor hydrologic condition (Hydrologic alteration index score > 0)	Good hydrologic condition (Hydrologic alteration index score = 0)
Poor biology (CSCI < 0.79)	***Flow Management***: Evaluate hydrologic alteration among other stressors. Determine relative importance of flow management for improving biological condition, relative to other stressors.	***Other Management/Causal Assessment***: Evaluate other stressors to determine cause of poor biology. Evaluation of flow management not recommended.
Good biology (CSCI > 0.79)	***Monitor***: Communities may be resilient to flow alteration. Continue to monitor for factors that may reduce resilience.	***Protect:*** Intact area. Target for preservation. Explore factors that may contribute to resilience or vulnerability.

4.2.7 Application to a regional survey

Sites sampled as part of an ongoing regional ambient monitoring program that uses a probabilistic survey design were used to estimate the extent of hydrologically altered streams, as well as the extent of the different management priority classes for the entire South Coast region. Sites were assessed based on major land use types (i.e., agricultural, urban, and open), as well as for major hydrologic stream classifications (Pyne et al., in review and previous chapter). Because sites were sampled under multiple surveys, weights were recalculated through post-stratification. These weights were used to estimate extent and magnitude using the Horvitz-Thompson estimator (Horvitz-Thompson 1952). Confidence intervals were based on local neighborhood variance estimators (Stevens and Olsen 2004). All calculations were conducted using the spsurvey package (Kincaid and Olsen 2013) in R (R Core Team 2016). Additional details about weight adjustments, land use classifications, and extent estimates are provided in Mazor (2015).

4.2.8 Comparison of the influence of hydrologic alteration and water chemistry on biological condition

We graphically evaluated relationships between hydrologic and chemical stressors and biology by plotting stressors against CSCI scores, and by constructing a linear model to predict CSCI scores from the index of hydrologic alteration, Chloride, total Nitrogen, and the ratio of observed to expected specific conductivity at the 124 sites where all data were available. In addition, relationships were explored qualitatively in a nonmetric multidimensional scaling (NMDS). An ordination was constructed using the metaMDS function in the vegan package in R (Oksanen et al. 2016 R Core Team 2016). Invertebrate data were processed as required to calculate the O/E component of the CSCI; specifically, taxa were aggregated to unambiguous operational taxonomic units, ambiguous data were tossed, and samples were standardized to 400-count samples. A 2-dimensional solution based on Bray-Curtis distance was calculated using default settings (apart from suppressing auto-transformation of thee data). Ordination axes were then used for correlation with hydrologic alteration and water chemistry variables using Spearman's rank correlation coefficient.

4.3 Results

4.3.1 Hydrologic alteration targets to support biological integrity

Hydrologic alteration based on most flow metrics was associated with a decline in biological index scores. In many cases, both increases and decreases in flow metrics were associated with biological degradation, and healthy conditions were most common where alteration was close to zero (e.g., Figure 4-3a). However, relationships were sometimes evident for just a single direction of alteration (increasing or decreasing; Figure 4-3b). Typically, this situation occurred where alteration along one gradient affected few sites, or affected them to a lesser degree than the other gradient. Within the data set, gradients of alteration were evident for nearly every metric, and all but one (i.e., fraction of years with no flow) had sufficient data for analysis (i.e., 30 or more sites exhibiting alteration along a single gradient).

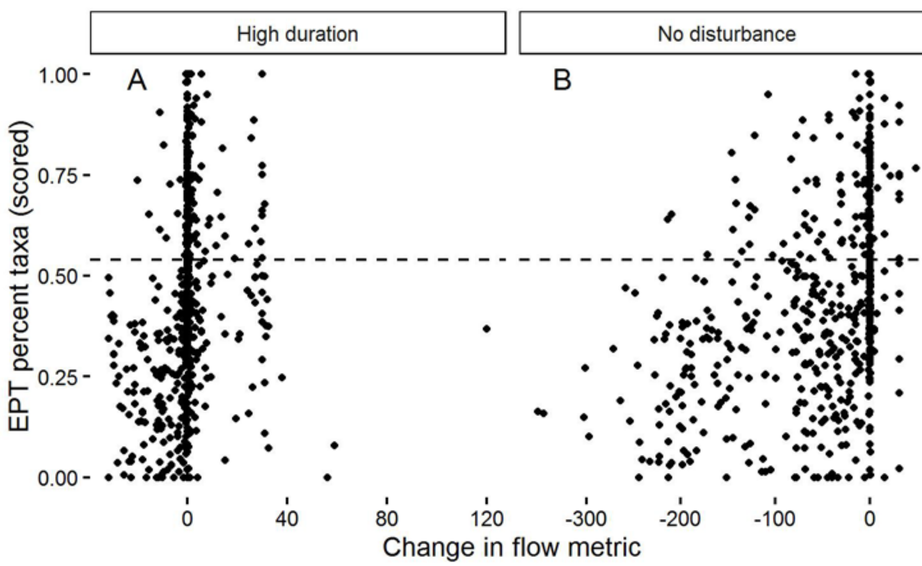

Figure 4-3. An example plot of biological change versus two measures of hydrological change. The dashed line is the threshold for identifying healthy versus degraded biological conditions. EPT: Ephemeroptera, Plecoptera, and Trichoptera. As described in Mazor et al. (2016), the biological metric is the difference between observed and predicted values (O/E ratio), transformed to a scale from 0 to 1.

Each biological endpoint was successfully modeled against nearly all of the flow metrics for all climatic conditions, with the exception of two variables (i.e., percent intolerant and shredder taxa). These endpoints rarely or never indicated poor condition (respectively), a consequence of the low mean and high variability of these metrics in the reference data set (Table 4-1). Although targets could be identified for multiple biological endpoints, the most conservative target was almost always associated with the MMI component of the CSCI. Selected relationships between flow metrics and the CSCI are shown in Figure 4-4.

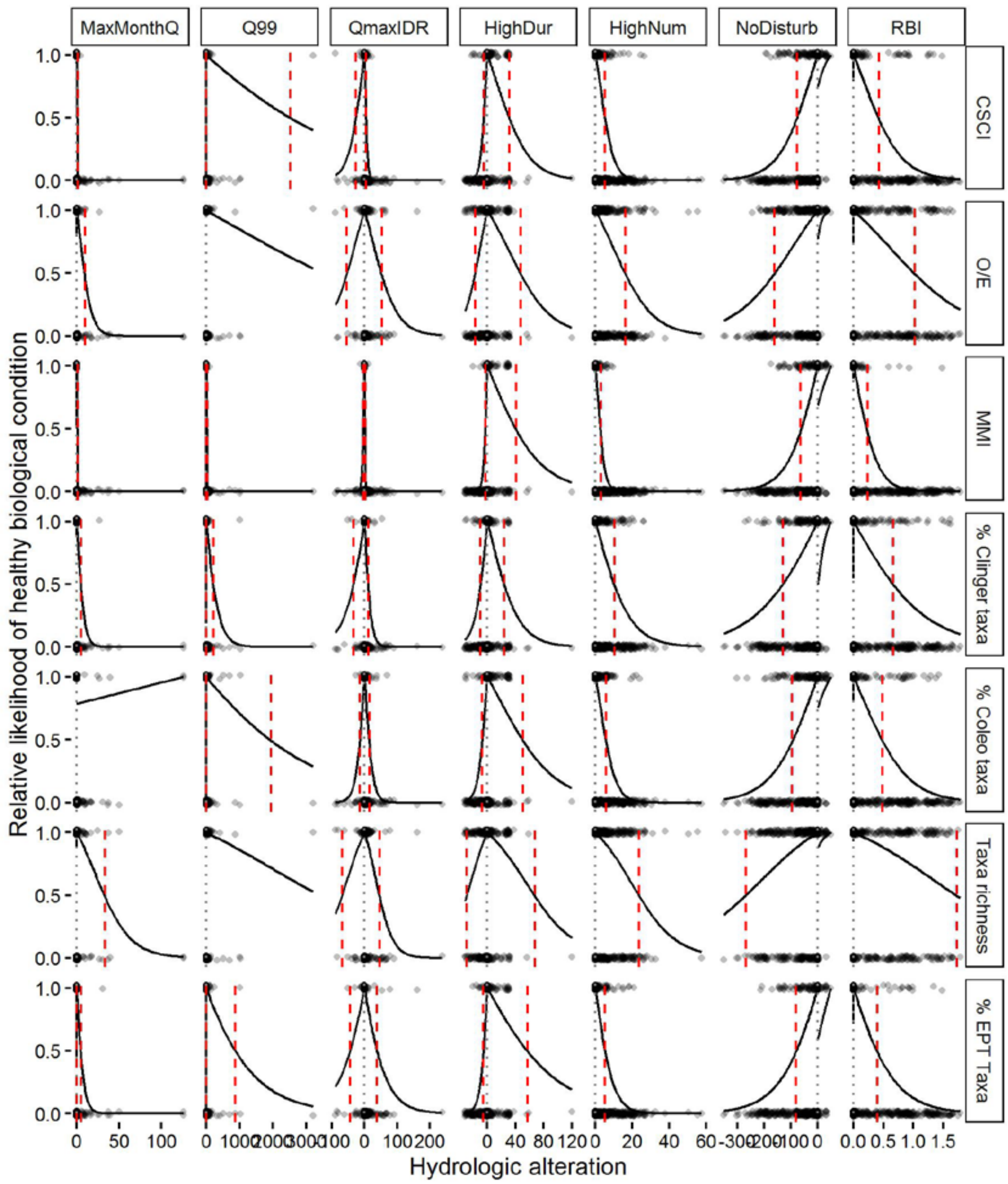

Figure 4-4. Relative likelihood of healthy biological conditions at different levels of hydrologic alteration for selected flow metrics. Points at the top of each panel represent sites in healthy biological condition, and points at the bottom of each panel represent sites in poor biological condition. Dotted vertical lines represent hydrologically unaltered conditions. Dashed vertical red lines represent targets where the likelihood is half the likelihood at unaltered conditions. Not shown: shredder taxa and % intolerant biological response variables. Coleo: Coleoptera.

Flow ecology relationships were stronger when based on all streams in the region than when based on streams from a single hydrologic class, as many class-specific relationships lost statistical significance or had too few data to analyze (Table 4-4). In only one case (i.e., the increasing gradient for Qmed in Class 1) was a target set for a class when we were unable to do so with the regional data set. When targets were successfully set for specific stream classes, some thresholds became stronger (i.e., more conservative) and others became weaker. Stronger metrics were more common for increasing gradients, and for streams in Class 4, whereas weaker targets were more common for decreasing gradients and for streams in Class 1. In a handful of cases, the targets for individual stream classes were identical to regional targets.

Table 4-4. Targets for selected hydrologic stream classes. Symbols indicate whether the target was more conservative (+), less conservative (-), or equal to (=) regional targets. ND: Insufficient data to analyze a target within a class. LS: No significant target within a class, but a target was set at the regional scale. NS: No significant target within a class, nor at the regional scale. GS: Target set within a class, but not at the regional scale. Blank cells indicate that the data were insufficient to set a target at the regional scale, and were therefore not analyzed for individual classes.

	Decreasing				Increasing		
Metric	**Class 1**	**Class 2**	**Class 4**		**Class 1**	**Class 2**	**Class 4**
Duration							
HighDur	ND	-	+		-	+	+
Hydroperiod					+	-	+
LowDur	ND	-	LS		ND	+	-
NoDisturb	-	-	+		ND	ND	ND
Per_LowFlow	-	+	+		LS	ND	ND
Frequency							
HighNum					-	=	+
MedianNoFlowDays	ND	+	LS				
Magnitude							
MaxMonthQ	ND	ND	ND		+	=	+
MinMonthQ	-	-	-		+	+	LS
Q01	-	-	=		ND	NS	ND

	Decreasing				Increasing		
Metric	**Class 1**	**Class 2**	**Class 4**		**Class 1**	**Class 2**	**Class 4**
Q05	ND	-	LS		LS	LS	LS
Q10	-	-	-		ND	NS	ND
Q25	-	-	-		NS	NS	NS
Q50	LS	-	-		+	-	+
Q75	LS	-	-		-	-	+
Q90	LS	-	LS		=	LS	+
Q95	LS	+	+		-	LS	+
Q99	LS	+	ND		-	+	=
Qmax					LS	+	+
Qmean					-	+	=
QmeanMEDIAN	+	LS	LS		+	+	LS
Qmed	LS	NS	LS		GS	NS	NS
Qmin	+	LS	LS		NS	NS	NS
Timing							
C_C	-	=	+		ND	ND	ND
C_CP	-	=	+		ND	ND	ND
C_M	ND	ND	ND		-	-	+
C_MP	ND	ND	ND		-	=	+
C_P	ND	+	+		ND	ND	+
MaxMonth	ND	ND	ND				
MinMonth	ND	-	-		ND	-	ND
Variability							

	Decreasing				Increasing		
Metric	**Class 1**	**Class 2**	**Class 4**		**Class 1**	**Class 2**	**Class 4**
QmaxIDR	ND	ND	ND		+	=	+
QmeanIDR	ND	ND	ND		+	+	=
QminIDR	LS	+	=		ND	ND	ND
RBI	ND	ND	ND		+	-	+
SFR	+	+	=		ND	ND	ND

4.3.2 Selection of hydrologic metrics and development of a hydrologic alteration index

The relative influence of flow metrics varied by metric class, but not by precipitation condition. Magnitude metrics (particularly those associated with high flows) and variability metrics showed the greatest influence on biological response variables. Certain metrics within the frequency and duration classes also had great influence. In contrast, timing metrics had relatively little influence over most response variables; the only exception being C_MP (Colwell's maximized contingency, a measure of the likelihood that the pattern of high- and low-flow events is repeated across years), which had large influence over the percent intolerant and taxonomic richness metrics (Figure 4-5).

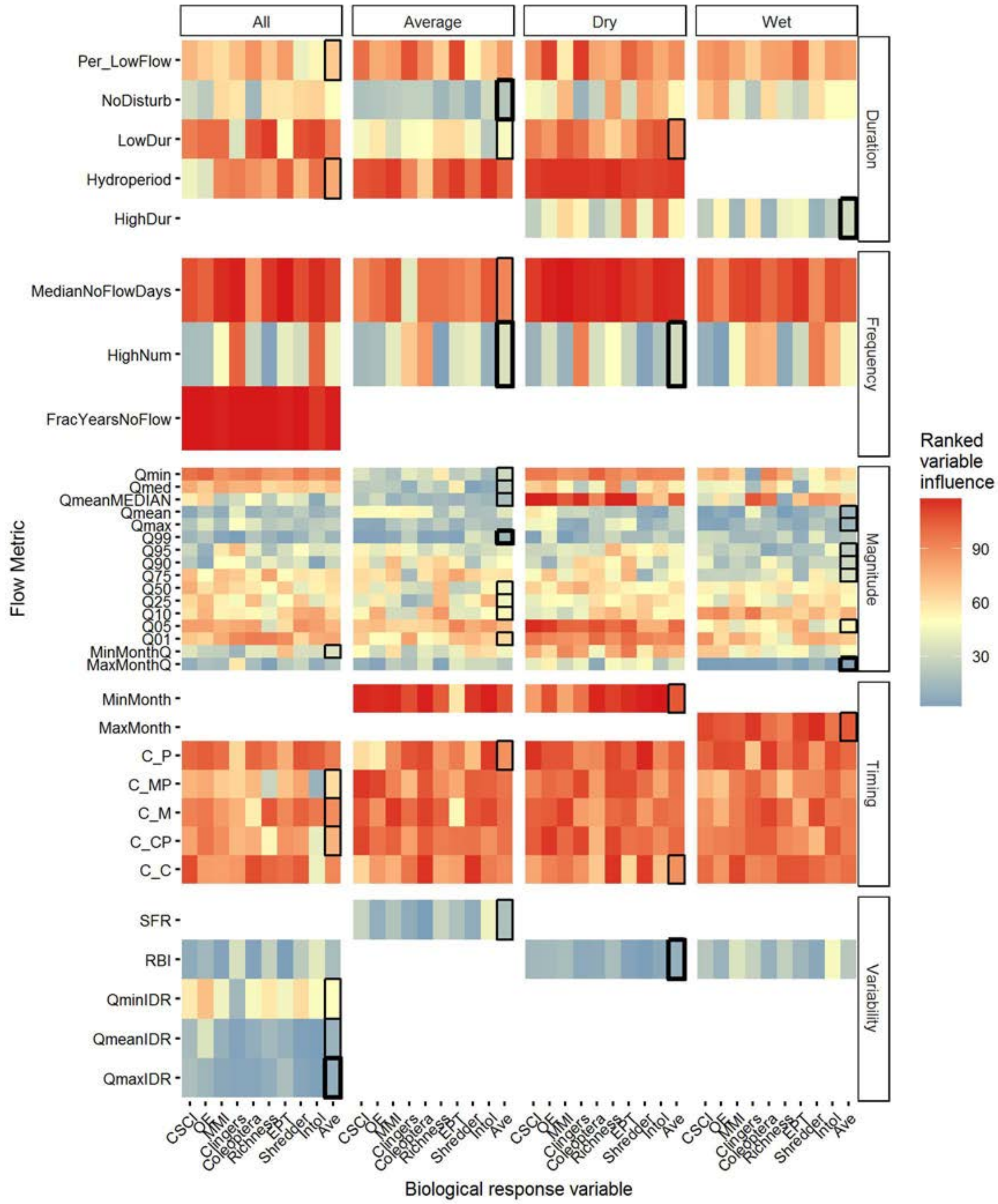

Figure 4-5. Ranked variable influence in boosted regression trees to predict biological response variables from metrics of flow alteration. Cell color indicates the ranked influence of the flow metric, and the "Ave" column indicates the average rank across the 9 biological response variables; blue cells are better ranked than red cells. White cells indicate flow metric-precipitation combinations that were not analyzed. Outlined cells indicate the best-ranked precipitation condition within a flow metric, and thick outlined cells indicate the metrics that were selected for inclusion in an index of flow alteration (i.e., up to two top-ranked metrics per class, and in the top half of metrics overall).

Based on the average ranked influence from the BRT analysis, seven flow metrics were selected for inclusion in an index of hydrologic alteration (Table 4-5). Metrics were prioritized based on the following criteria (Mazor et al. in review):

- Ability to differentiate reference sites and non-reference sites
- Strong relationship to biological condition based on boosted regression tree analysis and can produce a hypothesized ecological response
- Ability to be modeled under both current and reference conditions with a high level of confidence
- Amenability to management actions, with predictable responses to changes in flow conditions
- Representation of different components of the hydrograph (e.g. magnitude vs. duration)
- Minimal redundancy with other metrics

Table 4-5. Priority hydrologic metrics and associated thresholds used in the regional flow-ecology relationships. Metrics are grouped the hydrograph component they represent. Thresholds are expressed as the change in metric value (delta H) associated with poor biologic condition (CSCI <0.79). Thresholds can be based on increasing or decreasing flows. Metric effects on biology were typically strongest during either average, wet, or dry rainfall years, or all years combined (overall). NT= no threshold established.

Hydrograph Component	Metric	Metric Definition	Critical precipitation condition	Decreasing Threshold	Increasing Threshold
Duration	NoDisturb (days)	median annual longest number of consecutive days that flow is between the low and high flow threshold	Average	-64	NT
	HighDur (days/event)	median annual longest number of consecutive days that flow was greater than the high flow threshold	Wet	-3	24
Magnitude	MaxMonthQ (m3/s)	Maximum mean monthly streamflow	Wet	NT	1.5
	Q99 (m3/s)	streamflow exceeded 99% of the time	Wet	-0.01	32
Variability	RBI (unitless)	Richards-Baker index of stream flashiness	Dry	NT	0.25
	QmaxIDR (m3/s)	Interdecile range of flow	Overall	-5	2.5
Frequency	HighNum (events/year)	median annual number of events that flow was greater than high flow threshold	Dry	NT	3

Based on these criteria, we selected two duration (i.e., NoDisturb and HighDur), one frequency (i.e., HighNum), two magnitude (i.e., Q99 and MaxMonthQ) and two variability (i.e., RBI and QmaxIDR) metrics. Metrics based on high flows were favored over those associated with low flows, as they typically had greater influence in predicting biological responses. Because all timing metrics were among the bottom-ranked metrics, none were selected. Three of these metrics were based on average precipitation conditions, two on dry conditions, two on wet conditions, and one on overall conditions.

The overall hydrologic alteration index showed a negative relationship with all biological response variables (Figure 4-6). In general, the relationship was strongest when the hydrologic alteration score was below 5; the relationships for many response variables leveled off at higher levels of alteration, suggesting that benthic macroinvertebrate communities lack a capacity to respond to more severe levels of alteration or reach a saturation point above which there is no additional community response. The relationships were particularly striking for the shredder taxa and % intolerant metrics, despite the fact that these were rarely (for shredders) or never (for % intolerant) used to model responses to flow alteration.

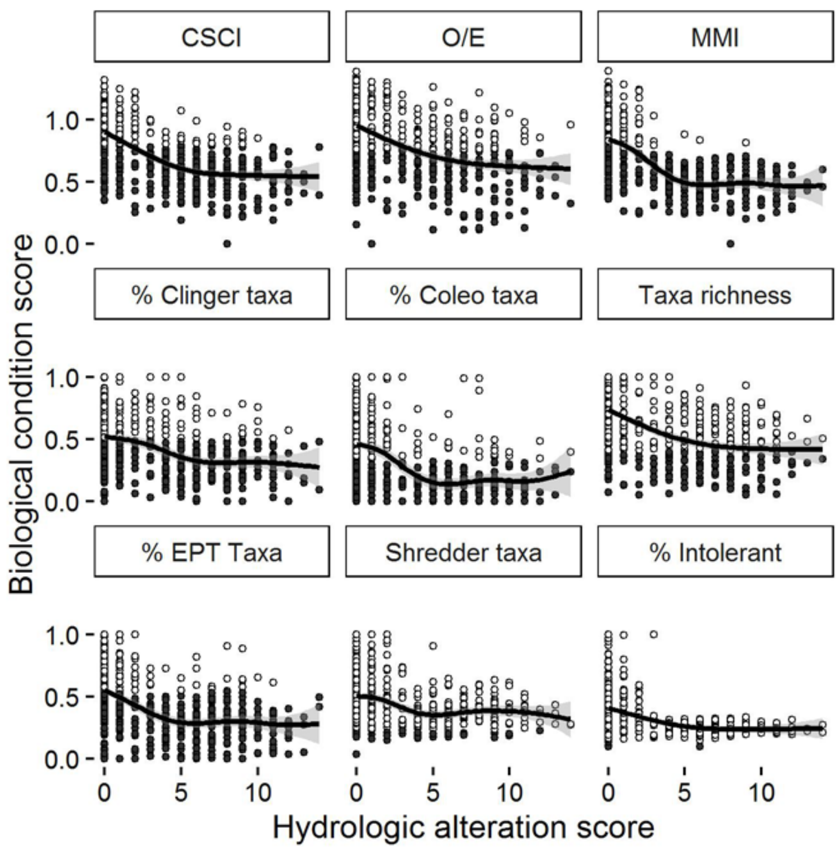

Figure 4-6. Biological responses to the index of hydrologic alteration. Sites in good biological condition are shown as white dots, and sites in poor biological condition are shown as black dots. The black line represents a smoothed fit from a general additive model, and the gray band represents its 95% confidence interval.

4.3.3 Classification and application to a regional survey

Application of the overall hydrologic alteration index to the regional probabilistic data set showed that about two-thirds of stream-kms were considered unaltered based on our criteria (59 to 73% of stream-kms, 95% confidence interval, n = 255 unaltered streams). Alteration was most extensive in urban streams (91%, n=177), followed by agricultural streams (80%, n = 44); alteration was limited to only 11% of stream-kms (n = 29) draining undeveloped catchments (Figure 4-7a). Among the seven hydrologic classes certain were more pervasively altered than others, although data were limited in a few classes. Among the three classes with the greatest extents within the region, class 1 was the least altered (14%, n = 14), while classes 2 and 4 were more extensively altered (37%, n = 91 and 40% of stream-kms, n = 81). Although very few sites were sampled in classes 5 and 7 (i.e., 6 and 10, respectively) and precision of estimates are therefore poor, they were found to be the most extensively altered (56% and 69%, respectively).

About half of the region's stream-kms (i.e., 52%, n = 183 sites) were designated for protection, as they had unaltered hydrology based on the index, and had good biological condition based on CSCI scores. These sites were predominantly located in mountainous areas (Figure 4-7b). Among streams with undeveloped catchments, about three-quarters (72% of stream-kms, n = 137) were in this category. A very small subset of streams (4% of stream-kms, n = 40) were in good biological condition, but showed evidence of hydrologic alteration, and were recommended for additional monitoring to track potential changes in condition. Although these streams were considered hydrologically altered, the alteration score was substantially and significantly lower than scores at sites where biological condition was poor (mean: 4.6 vs 6.7, Welch's two-sample t-statistic = 4.6, $p < 0.001$), suggesting that alteration may not be as severe in this group. About a quarter to a third of stream-kms (i.e., 30%, n = 227) were in poor biological condition and hydrologically altered, and evaluations of flow management to improve stream health are recommended for streams in this class; this class was particularly pervasive among urban (85%, n = 166) and agricultural (53%, n=31) streams. Finally, a small portion of the region (14%, n = 72) was in poor biological condition but unaltered hydrology. At these sites, stressors unrelated to flow alteration (such as degraded water quality or direct habitat modification) should be prioritized in causal assessments.

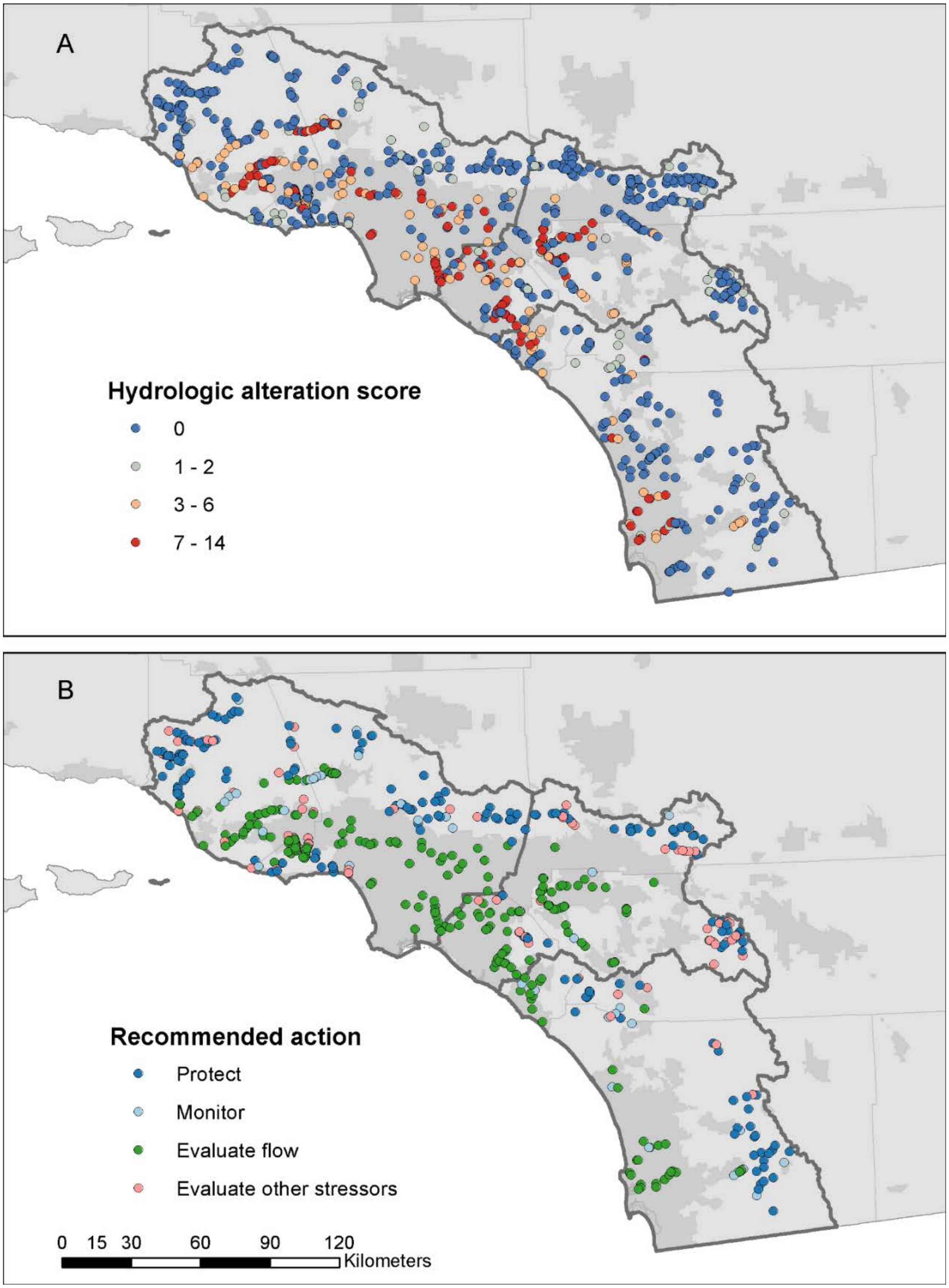

Figure 4-7. Hydrologic alteration scores (A) and recommended management actions (B) at sites in the region. Urban areas are represented as dark gray. Boundaries of major hydrologic regions are shown.

4.3.4 Comparison of the influence of hydrologic alteration and water chemistry alteration on biological condition

Hydrologic alteration was positively correlated with chemical alteration, although not statistically significant. Pearson's r^2 between the hydrologic alteration score was 0.22 for specific conductance, 0.27 for total Nitrogen, and 0.37 for Chloride. In general, sites in good biological condition were restricted to sites where these analytes were low, and where the hydrologic alteration score was low (Figure 4-8), supporting a role for both chemical and hydrological alteration in affecting biological condition. Linear model fit was good ($p<0.001$, adjusted R^2 0.41), but none of the water chemistry variables were statistically significant.

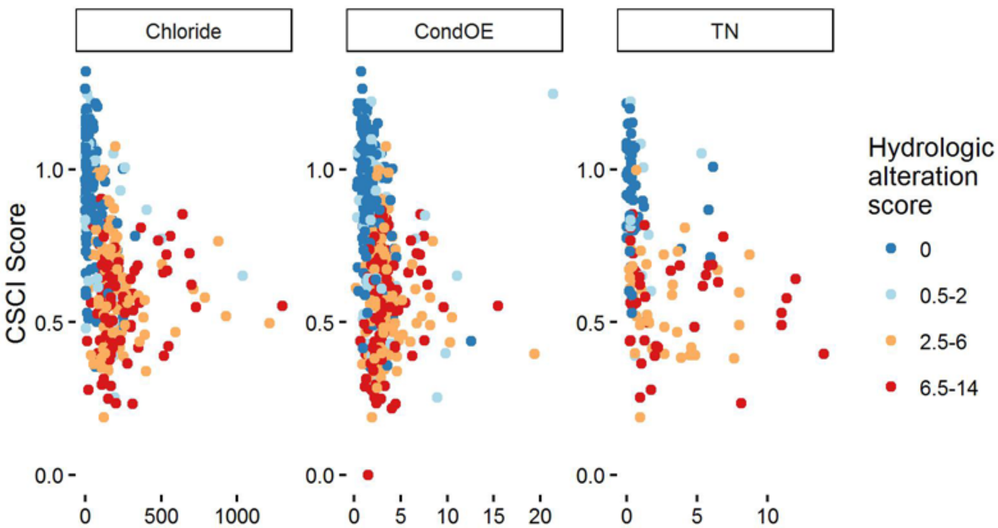

Figure 4-8. Relationship between CSCI scores, selected water chemistry analytes, and the hydrologic alteration score. Chloride and total nitrogen (TN) are in mg/L. CondOE is the ratio of the observed to expected specific conductance.

Ordination resulted in a moderately high-stress solution (stress = 0.257), yet a clear relationship with hydrologic alteration was evident (Figure 4-9a). Axis 1 was correlated with several measures of hydrologic alteration, including the overall hydrologic index (Spearman's rho: 0.41, Figure 4-9b). This axis was also positively correlated with several measures of water chemistry (e.g., chloride: 0.59; specific conductivity O/E: 0.46), and negatively correlated with the CSCI (-0.54). Flow metrics related to timing also showed strong relationships with this axis (e.g., Colwell's contingency (M): 0.41), despite having little influence on biological condition in boosted regression tree models. In contrast, several duration metrics had correlations with this axis that were weaker than might be expected (e.g., HighDur: -0.29); however, both increases and decreases in this metric were clustered on the right side of axis 1, which reduced the apparent strength of the relationship (Figure 4-9c). Many taxa considered to be sensitive to hydrologic alteration (e.g., Ephemeroptera, Plecoptera, and Trichoptera) were more common at the unaltered sites on the left side of Axis 1 (Figure 4-9d), whereas Odonata were more common at the altered sites on the right. Clingers in particular were more abundant at the sites with minimal flow alteration, and the % clinger taxa metric had a spearman rank correlation of -0.46 with Axis 1. Biological and environmental gradients along axis 2 were less obvious, although non-insects and Odonata were both more common at sites with high values on axis 2. The biological and environmental gradients associated with axis 2 were less clear, and very few hydrologic metrics had a stronger relationship with axis 2 than axis 1 (e.g., Colwell's predictability [P]).

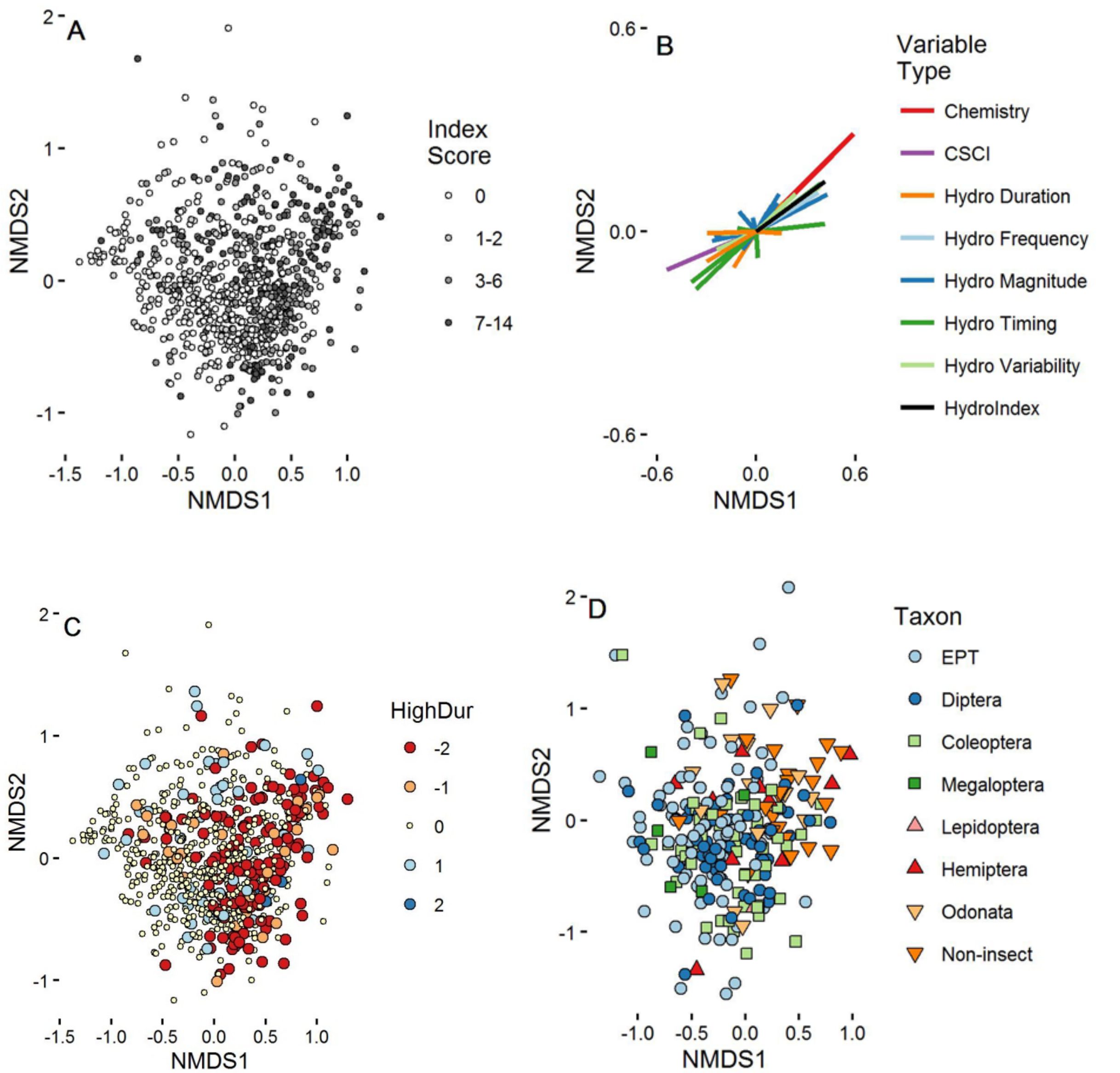

Figure 4-9. Nonmetric multidimensional scaling (NMDS) plots of sites in the study. A: Each point represents a single site. Darker colors indicate greater scores of the index of hydrologic alteration. B: Each vector represents a variable related to water chemistry, biology, or hydrologic alteration. Position of the endpoint of the vector indicates Spearman rank correlation metric with each axis. C: Same as panel A, except that symbol color and size indicates degree and alteration of the HighDur metric. D: Scores for each taxon in the data set. Symbol color and shape indicates taxonomic group. EPT: Ephemeroptera, Plecoptera, and Trichoptera.

4.4 Discussion

4.4.1 Regional targets create new options for managing hydrologic alteration

Regional flow targets associated with healthy biological communities provide an important tool to inform water management decisions and address effects of hydrologic alteration from local to regional scales. In particular, targets create a way for managers to anticipate the impacts of development, and to prioritize stream reaches or sub-watersheds for protection. Regionally-derived targets provide a robustness not possible with locally-derived thresholds because the flow-ecology relationships are based on large data sets. Although site-specific approaches may be appropriate for isolated, point-source impacts like dams, hydrologic alteration caused by urban runoff or increases in impervious surfaces can best be managed with a set of targets with regional applicability. By applying a regionally transferable ensemble of hydrologic models to a large bioassessment data set, we were able to model responses across a wide range of range of conditions, and derive targets that can be applied to sites throughout the region. Furthermore, because these targets are based on probabilistic relations in logistic regression models, managers can adjust the targets according to their tolerance for risk.

Planning and prioritization of management actions requires a regional understanding of hydrologic alteration. As demonstrated in this study, regional targets can be used for ambient assessments of hydrologic conditions (e.g., Figure 4-7a), and for rapid screening of hydrologic stressors associated with poor health (e.g., Figure 4-7b). Crucially, these targets allow the evaluation of best management practices, helping managers select the appropriate solutions for each site; for example, Stein et al. (in review) found that capture of peak storm flows could achieve targets more effectively than even great reductions in impervious surfaces in a small urban watershed in San Diego. Potentially, these targets could also be used to inform the development of policies that establish regional flow criteria, or to identify contributing stressors of biological impairment in the establishment of total maximum daily loads. However, these applications have not yet been evaluated, and we recommend further exploration through case studies and other investigations.

4.4.2 Integrative biological indicators of condition, like BMI, are essential targets for management

Although many studies of hydrologic alteration focus on impacts to fish and other vertebrates (e.g., Beecher et al. 2010, DePhilip and Moberg 2013, McManamay et al. 2013), they also recognize the importance of including additional community-level indicators of condition as a biological endpoint, such as benthic macroinvertebrate assemblages. Tools that focus on a single resource (e.g., minimum flows that support an endangered fish) are only useful where production of the resource is a goal. In arid regions like southern California, native fish fauna may be depauperate, and the majority of streams in the region may be fish-free from natural or anthropogenic reasons. Therefore, flow targets designed to protect fish are inappropriate for many streams in this region. In contrast, benthic macroinvertebrates are an appropriate management endpoint in nearly all streams (Bonada et al. 2006). Even in streams where fish are an appropriate endpoint, a single-species focus may not be integrative enough to adequately protect the health of streams. Managers should use flow targets designed to support integrative measures of biological integrity, such as benthic macroinvertebrates, in tandem with targets aimed at more specific resources of interest where needed. Expanding the manager's toolkit to include additional integrative biological indicators, such as benthic algae, may also have benefits.

Benthic macroinvertebrates are particularly good for assessing hydrologic alteration because they possess diverse life history traits that should respond strongly to alterations in flow regimes. For example, decreased low-flows may exacerbate stresses related to high temperature and low dissolved oxygen, favoring species with traits adapted to these conditions (e.g., respiratory pigments or air-breathing strategies). An increase in flashiness may frequently "reset" benthic macroinvertebrate communities through direct mortality; under this type of alteration, species that have resilient traits (e.g., strong aerial dispersal, propensity to drift, and rapid, multivoltine reproduction) may be favored. It is likely that many of the responses to flow alteration are mediated by habitat alteration. For example, lowland streams are often dominated by sands and fines, and increased runoff may incise the streambed to bedrock; in this scenario, burrowing mayflies may be extirpated not by the increase in Q99, but rather by the elimination of their preferred substrate (Walters 2011, Kath et al. 2016, Poff and Zimmerman 2010).

4.4.3 Relating hydrologic alteration to water quality and habitat degradation

As observed here and elsewhere (e.g., Buchanan et al. 2013, Carlisle et al. 2014), hydrologic alteration often co-occurs with water quality and habitat degradation, and rarely occurs in isolation from other stressors. Although this correlation creates challenges for understanding the root causes of biological degradation, it underscores one of the key benefits of a flow-focused approach to management: because hydrology is a master driver of stream condition, flow management may address multiple impacts at once. Management measures that reduce flow alteration are likely to also improve water quality, and may also lead to hydrologic regimes that generate the physical habitat conditions that support healthy biology. In some cases, however, poor water quality or habitat conditions may not respond to flow management. For example, concrete-lined channels (which comprise ~26% of stream-kms in the region, Mazor 2015) may never support the diversity of microhabitats that benthic macroinvertebrates require, even if natural hydrologic regimes are restored. Therefore, the value of flow management options should be evaluated in the context of how habitat and water quality conditions are likely to respond to improved hydrology.

4.4.4 Site-specific resilience can inform management, but requires further study

Although rare in our study (i.e., 40 sites), streams with altered hydrology but good biological condition were observed. While this discordance between hydrological and biological conditions could simply be due to statistical noise, it could also arise from differences in resilience to hydrologic alteration among watersheds. Several factors could contribute to resilience. At the catchment scale, some watersheds may be better able to absorb the impacts of dams, diversions, or increased imperviousness than others, dampening the impacts on the hydrograph. At the reach scale, certain types of habitats are more likely to tolerate more severe hydrologic alteration than others. For example, bedrock-dominated streams are less likely to erode in response to increased peak flows than streams dominated by fine substrates. At the organismal scale, certain life history traits confer a natural resilience to hydrologic alteration. These traits, such as rapid life-cycles and good dispersal ability, may be particularly important in regions with naturally high hydrologic variability, such as the arid Mediterranean climates found in California (Gasith and Resh 1999, Bonada et al. 2007). Although only a small number of sites were observed to be in good condition despite hydrologic alteration, it is possible that watershed-, reach-, and organismal-scale factors all contribute to the resilience observed at these sites.

4.4.5 Comparison with other studies reveal common themes

We found that metrics related to high flows were particularly useful in assessing biological responses to hydrologic alteration. For example, HighDur, HighNum, Q99, and MaxMonthQ were all selected for inclusion in the index, and similar metrics were highly ranked in BRT models. Similarly, Buchanan et al. (2013) observed strong responses to alterations the magnitude, duration, and frequency of high flow events. Although we successfully established targets for metrics related to low flows, they were less influential on biological response variables, and none were selected for inclusion in the hydrologic alteration index. There are a few explanations for this pattern, which contrasts with other studies (e.g., Bonada et al. 2007, Konrad et al. 2008, Kennen et al. 2010, Buchanan et al. 2013) that show that benthic macroinvertebrate communities are highly sensitive to changes in low-flow metrics. First, the local fauna may be well adapted to the pressures of low-flow conditions as a consequence of the region's arid, Mediterranean climate (Gasith and Resh 1999, Resh et al. 2013). Mazor et al. (2014) showed that an index of biotic integrity was robust to seasonal drought in southern California intermittent streams, underscoring the resilience of species in the region to low-flow conditions. Second, we were less successful in predicting low-flow metrics than high-flow metrics (Sengupta et al. in review), and the imprecision in these predictions may have caused us to underestimate their influence on benthic macroinvertebrates.

5.0 Conclusions and Next Steps

Using the general approach of ELOHA framework, we developed regional flow-ecology relationships based on healthy benthic macroinvertebrate communities. Following classification of California's wadeable streams into seven hydrologic classes, we focused the development of hydrologic targets and assessment of hydrologic alternations on the predominant stream classes in Southern California. Our analysis produced a set of ensemble hydrologic models that allow estimation of hydrologic alteration at any ungauged stream in the region based on 37 flow metrics under several climatic conditions (e.g. wet years, dry years, average years). We identified seven priority flow metrics that are most associated with biological effects, representing a broad range of hydrologic properties (e.g. flow magnitude, duration, frequency), that can be readily estimated, and are amenable to management actions. Using these seven flow metrics, we were able to assess the extent of hydrologic alteration in the region and categorize streams into four major management classes.

Three key factors contributed to our success in setting regional thresholds for flow alteration: 1) an ensemble of regionally transferable hydrologic models, 2) a large bioassessment data set from regional surveys, and 3) a widely accepted index of stream health (i.e., the CSCI) that is minimally influenced by natural co-factors. The ensemble of 26 models allows estimation of hydrologic alteration at any ungauged site, provided that adequate precipitation data are available (Sengupta et al. in review). Expanding the ensemble with additional models would likely improve performance at stream-types that are currently underrepresented (e.g., small high-elevation watersheds), and allow expansion outside the South Coast region. The bioassessment data, most importantly, represented a wide range of conditions, from minimally disturbed reference sites (Stoddard et al. 2006, Ode et al. 2016) to sites where hydrologic alteration was severe. Although only a small subset of these sites were co-located with stream gauges (i.e., 133, of which 20 were reference), the transferability of the hydrologic models allowed us to take full advantage of the data set and the stressor gradients it represents. Because the index of stream health is based on predictive models that set site-specific biological expectations in environmental context (e.g., climate, watershed area, geology), scores have the same meaning at all sites in the study area. Thus, the index could be used to characterize changes in biological condition for multiple river-types, without the classification steps ordinarily required in the ELOHA approach (Poff et al. 2010). Indeed, this index, combined with the regional transferability of the hydrologic models, likely accounts for the small differences in targets we observed among stream classes (Table 4-4). Generating targets for other biological endpoints, such as fish, benthic algae, or riparian vegetation, may require additional analyses to control the influence of natural gradients if predictive models or indices are unavailable.

The transferability of the hydrologic models is enhanced by their simplicity, but this simplicity brings with it a few downsides. For example, in our study, historic conditions were simulated by altering mechanistic model parameters related to watershed imperviousness, and it could be argued that our models show responses to land use rather than to hydrologic alteration. However, by translating changes in imperviousness into biologically relevant targets for flow alteration, we've created tools that can be used in a variety of applications discussed above, including specification of stormwater control measures that maintain and reestablish target streamflows. In contrast, targets for imperviousness could not be used for most management applications, like causal assessment or comparing restoration options. More complex models (such as ParFlow.CLM; Bhaskar et al. 2015) can incorporate many more factors (such as coupled groundwater-surface water-land surface interactions), and thereby provide more realistic estimates of current and historic flow conditions, but this complexity makes them impractical for regional applications at hundreds of sites, and so were inappropriate for our study. Complex models could be used at sites of interest, in conjunction with the targets derived from simpler models, to evaluate management decisions. Therefore, the simplicity of the hydrologic models does not detract from their utility.

The flow-ecology tools and targets can be used in a variety of ways, several of which have been demonstrated in the San Diego River Watershed (Stein et al. 2016). In addition to mapping hydrologic alteration and identified unmodified areas for protection, the flow ecology tools can be used to establish management or compliance targets relative to flow, support causal assessments to determine whether flow alteration is a dominant stressor affecting biological condition, and to inform development of management actions aimed at mitigating or remediating the effects of hydrologic alteration.

Future efforts should expand to include stream classes from regions outside southern California. These efforts should also focus on enhancing hydrologic modeling to more fully assess hydrologic alteration due to groundwater infiltration or withdrawal or stream diversion. Hydraulic modeling could be used to more directly assess the effect of physical habitat alteration on biological condition, relative to changes in flow. Finally, our analysis could be repeated for other biological indicators of interest, such as algae, fish, or amphibians in order to provide a broader suite of tools for assessing condition, evaluating hydrologic stress, and setting management targets.

6.0 LITERATURE CITED

Anderson, M. J. 2001. Permutation tests for univariate or multivariate analysis of variance and regression. *Canadian journal of fisheries and aquatic sciences, 58*(3), 626-639.

Anderson, M. J., & Walsh, D. C. 2013. PERMANOVA, ANOSIM, and the Mantel test in the face of heterogeneous dispersions: what null hypothesis are you testing?. *Ecological Monographs, 83*(4), 557-574.

Baker, D.B., R.P. Richards, T.T. Loftus, and J.W. Kramer. 2004. A new flashiness index: Characteristics and applications to Midwestern Rivers and streams. JAWRA Journal of the American Water Resources Association, 40: 503–522. doi:10.1111/j.1752-

Beecher, H.A., B.A. Caldwell, S.B. DeMond, D. Seiler, S.N. Boessow. 2010. An Empirical Assessment of PHABSIM Using Long-Term Monitoring of Coho Salmon Smolt Production in Bingham Creek, Washington, North American Journal of Fisheries Management, 30: 6, 1529 — 1543: DOI: 10.1577/M10-020.1

Belmar O., J. Velasco and F. Martinez-Capel. 2011. Hydrological classification of natural flow regimes to support environmental flow assessments in intensively regulated Mediterranean rivers, Segura River Basin (Spain). Environmental Management, Supplement 47: 992-1004. 10.1007/s00267-011-9661-0.

Besaw, L.E., D.M. Rizzo, P.R. Bierman and W.R. Hackett. 2010. Advances in ungauged streamflow prediction using artificial neural networks. *Journal of Hydrology, 386*(1), pp.27-37.

Bhaskar, A.S., C. Welty, R.M. Maxwell, and A.J. Miller. 2015. Untangling the effects of urban development on subsurface storage in Baltimore. Water Resources Research, 51(2), 1158-1181

Bonada, N., N. Prat, V.H. Resh, and B. Statzner. 2006. Developments in aquatic insect biomonitoring: A comparative analysis of recent approaches. Annual Review of Entomology 51: 495-523.

Bonada, N., M. Rieradevall and N. Pratt. 2007. Macroinvertebrate community structure and biological traits related to flow permanence in a Mediterranean river network. Hydrobiologia 589: 91-106.

Boulton, A.J. 2014. Conservation of ephemeral streams and their ecosystem services: What are we missing? Aquatic Conservation: Marine and Freshwater Ecosystems 24: 733-738.

Brabec, E., S. Schulte and P.L. Richards. 2002. Impervious surfaces and water quality: a review of current literature and its implications for watershed planning. *Journal of planning literature, 16*(4), pp.499-514.

Breiman L. 2001. Random forests. Machine Learning 45: 5-32.

Brown, S. C., R.E. Lester, V.L. Versace, J. Fawcett, and L, Laurenson. 2014. Hydrologic landscape regionalisation using deductive classification and random forests. Plos One, 9, e112856.

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, USA. Freshwater Biology 58: 2632-2647. 10.1111/fwb.12240.

Bunn, S.E. and A.H. Arthington. 2002, Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity: Environmental Management, v. 30, no. 4, p. 492–507.

Cafri G. 2013. Predictors of students switching out of stem majors. San Diego State University, San Diego, California, U.S.A.

Carlisle D.M., J. Falcone, D.M. Wolock, M.R. Meador and R.H. Norris. 2010. Predicting the natural flow regime: models for assessing hydrological alteration in streams. River Research and Applications 26: 118-136. 10.1002/rra.1247.

Carlisle, D.M., S.M. Nelson, and K. Eng. 2014. Macroinvertebrate Community Condition Associated with the Severity of Streamflow Alteration. River Research and Applications 30: 29-39. Doi 10.1002/Rra.2626.

Carlisle, D.M., D.M. Wolock and M. Weiczorek. In Review. Estimating Natural Monthly Streamflows in California and their Likelihood of Anthropogenic Modification. USGS, Open-file report.

Colwell, R.K. 1974. Predictability, constancy, and contingency of periodic phenomena. Ecology 55(5): 1148-53.

Datry, T. 2012. Benthic and hyporheic invertebrate assemblages along a flow intermittence gradient: effects of duration of dry events. Freshwater Biology 57: 563-574. 10.1111/j.1365-2427.2011.02725.x.

DeGasperi, C.L., H.B. Berge, K.R. Whiting, J.J. Burkey, J.L. Cassin and R.R. Fuerstenberg. 2009. Linking Hydrologic Alteration to Biological Impairment in Urbanizing Streams of the Puget Lowland, Washington, USA. Journal of the American Water Resources Association (JAWRA) 45(2):512-533. DOI: 10.1111/j.1752-1688.2009.00306.x

DePhilip, M. and T. Moberg. 2010. Ecosystem flow recommendations for the Susquehanna River Basin. The Nature Conservancy. Harrisburg, PA.

DePhilip, M. and T. Moberg. 2013. Ecosystem flow recommendations for the Upper Ohio River basin in western Pennsylvania. The Nature Conservancy. Harrisburg, PA.

Elith, J., J.R. Leathwick and T. Hastie. 2008. A working guide to boosted regression trees. Journal of Animal Ecology 77: 802–813.

Eng, K., D.M. Wolock and D.M. Carlisle. 2013. River flow changes related to land and water management practices across the conterminous United States. Science of the Total Environment 463-464: 412-422.

Falcone, J. 2012. GAGES-II: Geospatial Attributes of Gages for Evaluating Streamflow. U.S. Geological Survey, Reston, Virginia, U.S.A.

Falcone. J.A., D.M. Carlisle, D.M. Wolock, and M.R. Meador. 2010. GAGES: A stream gage database for evaluating natural and altered flow conditions in the conterminous United States. Ecology 91: 621-621. 10.1890/09-0889.1.

Friedman, J.H. 2001. Greedy function approximation: A gradient boosting machine. The Annals of Statistics 29: 1189–1232.

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 and Systematics 30:51–81.

Gesch, D., M. Oimoen, S. Greenlee, C, Nelson, M. Steuck and D. Tyler. 2002. The national elevation dataset. *Photogrammetric engineering and remote sensing*, *68*(1), pp.5-32.

Grantham, T.E., J.H. Viers and P.B. Moyle. 2014. Systematic screening of dams for environmental flow assessment and implementation. BioScience 64: 1006-1018

Hanak, E. and J.R. Lund, J.R. 2012. Adapting California’s water management to climate change. *Climatic Change*, *111*(1), pp.17-44.

Hering, J.G., T.D. Waite, R.G. Luthy, J.E. Drewes and D.L. Sedlak. 2013. A changing framework for urban water systems. *Environmental science & technology*, *47*(19), pp.10721-10726.

Hollis, G.E. 1975. The effect of urbanization on floods of different recurrence interval. Water Resources Research 11: 431-435. 10.1029/WR011i003p00431.

Holmes, M.G.R., A.R. Young, A. Gustard and R. Grew. 2002. A region of influence approach to predicting flow duration curves within ungauged catchments. *Hydrology and Earth System Sciences Discussions*, *6*(4), pp.721-731.

Horvitz, D.G. and D.J. Thompson. 1952. A generalization of sampling without replacement from a finite universe. Journal of the American Statistical Association 47:663-685.

Jain S.K. & Sudheer K.P. (2008) Fitting of hydrologic models: a close look at the Nash–Sutcliffe index. Journal of Hydrologic Engineering, **13**(10), 981-986.

Karr, J. R. and E.W. Chu. 2000. Sustaining living rivers. Hydrobiologia 422/423: 1-14.

Kath J., E. Harrison, B.J. Kefford, L. Moore, P.J. Wood, R.B. Schäfer, F. Dyer. 2016. Looking beneath the surface: using hydrogeology and traits to explain flow variability effects on stream macroinvertebrates. Ecohydrology (online). DOI: 10.1002/eco.1741

Kennen, J. G., Riva-Murray, K., & Beaulieu, K. M. 2010. Determining hydrologic factors that influence stream macroinvertebrate assemblages in the northeastern US. *Ecohydrology*, *3*(1), 88-106.

Kennard, M.J., B.J. Pusey, J.D. Olden, S.J. MacKay, J.L. Stein, and N. Marsh. 2010. Classification of natural flow regimes in Australia to support environmental flow management. Freshwater Biology 55: 171-193. 10.1111/j.1365-2427.2009.02307.x.

Kincaid, T.M. and A.R. Olsen. 2013. Spsurvey: Spatial survey design and analysis. R package version 2.6. R Foundation for Statistical Computing. Vienna, Austria.

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. Predicting daily flows in ungauged catchments: model regionalization from catchment descriptors at the Coweeta

Hydrologic Laboratory, North Carolina. *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, CP, A.M.D 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: 1983-1998. 10.1111/j.1365-2427.2008.02024.x.

Konrad, C.P., D.B. Booth and S.J. Burges. 2005. Effects of urban development in the Puget Lowland, Washington, on interannual streamflow patterns: consequences for channel form and streambed disturbance. *Water Resources Research*, *41*(7).

Kuhn, M., J. Wing, S. Weston, A. Williams, C. Keefer and A. Engelhardt. 2012. Caret: Classification and Regression Training. R package version 5.15-045.

Lane, B.A., S. Sandoval-Solis and H.E. Dahlke. In Press. Hydrologic classification of natural flow regimes for regional environmental flows applications. Journal of the American Water Resources Association

Liaw, A. and M. Wiener. 2002. Classification and regression by randomForest. *R news*, *2*(3), pp.18-22.

Liermann, C.A.R., J.D. Olden, T.J. Beechie, M.J. Kennard, P.B. Skidmore, C.P. Konrad and H. Imaki H. 2012. Hydrogeomorphic Classification of Washington State Rivers to Support

Emerging Environmental Flow Management Strategies. River Research and Applications 28: 1340-1358. Doi 10.1002/Rra.1541.

Maechler M., Rousseeuw P., Struyf A., Hubert M. & Hornik K. (2015) Cluster: cluster analysis basics and extensions. *R Package version 2.0. 1*. **1**(2), 56

Mazor R.D., Stein E.D., Ode P.R. & Schiff K. (2014) Integrating intermittent streams into watershed assessments: applicability of an index of biotic integrity. *Freshwater Science*, **33**(2), 459-474.

Mazor, R.D. 2015. Bioassessment of Streams in Southern California: A Report on the First Five Years of the Stormwater Monitoring Coalition's Regional Stream Survey. Southern California Coastal Water Research Project Technical Report #844

Mazor, R.D., A.C. Rehn, P.R. Ode, M. Engeln, K.C. Schiff, E.D. Stein, D.J. Gillett, D.B. Herbst, C.P. Hawkins. 2016. Bioassessment in complex environments: designing an index for consistent meaning in different settings. Freshwater Science. 35(1):249–271. DOI: 10.1086/684130

Mazor, R.D., J. May, A. Sengupta, K. McCune, E.D. Stein. (In review). Setting hydrologic targets to support biological integrity. Freshwater Biology

Mazor, R.D., K. Schiff, K. Ritter, A. Rehn, P. Ode. 2010. Bioassessment tools in novel habitats: An evaluation of indices and sampling methods in low-gradient streams in California. Environmental Monitoring and Assessment 167:91-104.

McClain, M.E., A.L. Subalusky, E.P. Anderson, S.B. Dessu, A.M. Melesse, P.M. Ndomba, J.O.D. Mtamba, R.A. Tamatamah and C. Mligo. 2014. Comparing flow regime, channel hydraulics and biological communities to infer flow–ecology relationships in the Mara River of Kenya and Tanzania. Hydrological Sciences Journal, 59 (3–4), 801–819.

McCune, B. and J.B. Grace. 2002. Analysis of Ecological Communities. MjM Software Design, Gleneden Beach, Oregon.

McManamay, R.A., D.J. Orth, C.A. Dolloff and D.C. Mathews. 2013. Application of the ELOHA Framework to Regulated Rivers in the Upper Tennessee River Basin: A Case Study. Environmental Management 51:1210–1235: DOI 10.1007/s00267-013-0055-3.

Miller, S. W., Wooster, D., & Li, J. (2007). Resistance and resilience of macroinvertebrates to irrigation water withdrawals. *Freshwater Biology*, *52*(12), 2494-2510.

Moretti, G. and A. Montanari. 2008. Inferring the flood frequency distribution for an ungauged basin using a spatially distributed rainfall-runoff model. *Hydrology and Earth System Sciences*, *12*(4), pp.1141-1152.

Morley, S.A. and J.R. Karr. 2002. Assessing and Restoring the Health of Urban Streams in the Puget Sound Basin. Conservation Biology 16(6):1498–1509.

Naiman, R.J., S.E. Bunn, C. Nilsson, G.E. Petts, G. Pinnay, L.C. Thompson. 2002. Legitimizing fluvial ecosystems as users of water—an overview: Environmental Management, v. 30, no. 4, p. 455–467.

Nezlin, N.P. and E.D. Stein. 2005. Spatial and temporal patterns of remotely-sensed and field-measured rainfall in southern California. *Remote Sensing of Environment*, *96*(2), pp.228-245.

Novak, R., J.G. Kennen, R.W. Abele, C.F. Baschon, D.M. Carlisle, L. Dlugolecki., J.E. Flotermersch, P. Ford, J. Fowler, R. Galer, L.P. Gordon, S.N. Hansen, B. Herbold, T.E. Johnson,

J.M. Johnston, C.P. Konrad, B. Leamond and P.W. Seelbach. 2015. Draft: EPA-USGS Technical Report: Protecting Aquatic Life from Effects of Hydrologic Alteration: U.S. Geological Survey Scientific Investigations Report 2015–5160, U.S. Environmental Protection Agency EPA Report 822-P-15-002, XX p., http://pubs.usgs.gov/sir/2015/5160/ and http://www2.epa.gov/wqc/aquatic-life-ambient-water-quality-criteria

Ode, P.R., T.M. Kincaid, T. Fleming and A.C. Rehn. 2011. Ecological Condition Assessments of California's Perennial Wadeable Streams: Highlights from the Surface Water Ambient Monitoring Program's Perennial Streams Assessment (PSA) (2000-2007). A collaboration between the State Water Resources Control Board's Non-Point Source Pollution Control Program (NPS Program), Surface Water Ambient Monitoring Program (SWAMP), California Department of Fish and Game Aquatic Bioassessment Laboratory, and the U.S. Environmental Protection Agency.

Ode, P.R. 2007. Standard operating procedures for collecting benthic macroinvertebrate samples and associated physical and chemical data for ambient bioassessment in California. Surface Water Ambient Monitoring Program. Sacramento, CA.

Ode, P.R., A.C. Rehn, R.D. Mazor, K.C. Schiff, E.D. Stein, J.T. May, L.R. Brown, D.B. Herbst, D. Gillett, K. Lunde, and C.P. Hawkins. 2016. Evaluating the adequacy of a reference-site pool for ecological assessment in environmentally complex regions. Freshwater Science 35: 237-248.

Oksanen, J., F.G. Blanchet, M. Friendly, R. Kindt, P. Legendre, P.R. Minchin, R.B. O'Hara, G.L. Simpson, P. Solymos, M.H.M. Stevens, E. Szoecs and H. Wagner. 2016. vegan: Community Ecology Package. R. Package version 2.4-1.

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: 503-518. Doi 10.1002/Eco.251.

Olson, J.R. and C.P. Hawkins. 2012. Predicting natural base-flow stream water chemistry in the western United States. Water Resources Research 48: 10.1029/2011WR011088.

Parajka, J., A. Viglione, M. Rogger, J.L. Salinas, M. Sivapalan and G. Blöschl. 2013. Comparative assessment of predictions in ungauged basins–Part 1: Runoff-hydrograph studies. *Hydrology and Earth System Sciences*, *17*(5), pp.1783-1795.

Parasiewicz, P. 2009. Habitat Time Series Analysis to Define Flow Augmentation Strategy for the Quinebaug River, Connecticut and Massachusetts, USA. River Research and Applications. Published online in Wiley InterScience (www.interscience.wiley.com) DOI: 10.1002/rra.1066.

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: 1805-1818.

Poff. N.L., B.D. Richter, A.H. Arthington, S.E. Bunn, R.J. Naiman, E. Kendy, M. Acreman, C. Apse, 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: 147-170. 10.1111/j.1365-2427.2009.02204.x.

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: 71-91.

Poff, N. LeRoy, J. David Allan, Margaret A. Palmer, David D. Hart, Brian D. Richter, Angela H. Arthington, Kevin H. Rogers, Judy L. Meyer, and Jack A. Stanford. 2003. "River flows and water wars: emerging science for environmental decision making." *Frontiers in Ecology and the Environment* 1, no. 6: 298-306.

Poff, N.L. and J.K. Zimmerman. 2010. Ecological responses to altered flow regimes: a literature review to inform the science and management of environmental flows. *Freshwater Biology*, *55*(1), pp.194-205.

Poff, N.L. 2009. Managing for variability to sustain freshwater ecosystems.*Journal of Water Resources Planning and Management*, *135*(1), pp.1-4.

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), pp.769-784.

Poff, N.L., B. Richter, A.H. Arthington, S.E. Bunn, R.J. Naiman, C. Apse, E. Kendy, A.T. Warner, R. Tharme, B.P. Bledsoe, D. Merritt, R.B. Jacobson, M. Freeman, K. Rogers, J. Henriksen, J. Olden, J. O'Keeffe and M. Acreman. 2010. The ecological limits of hydrologic alteration: A framework for developing regional environmental flow standards. Freshwater Biology 55, 147–170.

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), pp.264-285.

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—a paradigm for river conservation and restoration: Bioscience, v. 47, no. 11, p. 769–784.

Poff, N.L., Olden, J.D., Merritt, D.M. and Pepin, D.M., 2007. Homogenization of regional river dynamics by dams and global biodiversity implications. *Proceedings of the National Academy of Sciences*, *104*(14), pp.5732-5737.

Poff, N.L., M.I. Pyne, B.P. Bledsoe, C.C. Cuhaciyan and D.M. Carlisle. 2010. Developing linkages between species traits and multiscaled environmental variation to explore vulnerability of stream benthic communities to climate change. *Journal of the North American Benthological Society*, *29*(4), pp.1441-1458.

Post, D.A. and A.J. Jakeman. 1999. Predicting the daily streamflow of ungauged catchments in SE Australia by regionalising the parameters of a lumped conceptual rainfall-runoff model. *Ecological Modelling*, *123*(2), pp.91-104.

Price, K., S.T. Purucker, S.R. Kraemer and J.E. Babendreier. 2012. Tradeoffs among watershed model calibration targets for parameter estimation. *Water Resources Research*, *48*(10).

Pringle, C. M., Freeman, M. C., & Freeman, B. J. 2000. Regional Effects of Hydrologic Alterations on Riverine Macrobiota in the New World: Tropical–Temperate Comparisons: The massive scope of large dams and other hydrologic modifications in the temperate New World has resulted in distinct regional trends of biotic impoverishment. While neotropical rivers have fewer dams and limited data upon which to make regional generalizations, they are ecologically vulnerable to increasing hydropower development and biotic patterns are emerging. *BioScience*, *50*(9), 807-823.

Puckridge, J.T., F. Sheldon, K.F. Walker and A.J. Boulton. 1998. Flow variability and the ecology of large rivers. *Marine and freshwater research*,*49*(1), pp.55-72.

Pyne, M.I., D.M. Carlisle, C.P. Konrad, E.D. Stein. (in press) Classification of California streams using combined deductive and inductive approaches: setting the foundation for analysis of hydrologic alteration. Ecohydrology (in press).

R Core Team. 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Version 3.2.4. Vienna, Austria.

Raymond, P. A., Neung-Hwan Oh, R. E. Turner, and W. Broussard. 2008."Anthropogenically enhanced fluxes of water and carbon from the Mississippi River." *Nature* 451, no. 7177: 449-452.

Reed, J.C., and C.A. Bush. 2005. Generalized geologic map of the conterminous United States. 2nd Edition. National Atlas. U.S. Geological Survey, Denver, Colorado, USA.

Resh, V.H., L.A. Bêche, J.E. Lawrence, R.D. Mazor, E.P. McElravy, A.P. O'Dowd, D. Rudnick and S. Carlson. 2013. Long-term population and community patterns of benthic macroinvertebrates and fishes in Northern California Mediterranean-climate streams. Hydrobiologia 719: 93. doi:10.1007/s10750-012-1373-9.

Richards, A.B. and D.C. Rogers. 2006. List of Freshwater Macroinvertebrate Taxa from California and Adjacent States Including Standard Taxonomic Effort Levels. Southwest Association of Freshwater Invertebrate Taxonomists. Chico, CA.

Richter B., Baumgartner J., Wigington R. & Braun D. (1997) How much water does a river need? *Freshwater biology*, **37**(1), 231-249.

Richter, B. D., Mathews, R., Harrison, D. L., & Wigington, R. (2003). Ecologically sustainable water management: managing river flows for ecological integrity. *Ecological applications*, *13*(1), 206-224.

Richter B.D., Warner A.T., Meyer J.L. & Lutz K. (2006) A collaborative and adaptive process for developing environmental flow recommendations. *River Research and Applications*, 22(3), 297-318.

Richter, B.D., M.M. Davis, C. Apse, C. Konrad. 2011. A presumptive standard for environmental flow protection. River Research and Applications 28(8):1312-1321. DOI: 0.1002/rra.1511

Ridgeway, G. 2015. gbm: Generalized Boosted Regression Models. R package version 2.1.1. https://CRAN.R-project.org/package=gbm

Rolls R.J. and A.H. Arthington. 2014. How do low magnitudes of hydrologic alteration impact riverine fish populations and assemblage characteristics? Ecological Indicators 39: 179-188. http://dx.doi.org/10.1016/j.ecolind.2013.12.017.

Rosenberg, D.M., P. McCully and C.M. Pringle. 2000. Global-scale environmental effects of hydrological alterations: introduction. *BioScience*,*50* (9), pp.746-751.

Samaniego, L., R. Kumar and S. Attinger. 2010. Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. *Water Resources Research*, *46*(5).

Sanborn, S.C. and B.P. Bledsoe. 2006. Predicting streamflow regime metrics for ungauged streams in Colorado, Washington, and Oregon. *Journal of Hydrology*, *325*(1), pp.241-261.

Schriever, T.A., M.T. Bogan, K.S. Boersma, M. Canedo-Arguelles, K.L. Jaeger, J.D. Olden and D.A. Lytle. 2015. Hydrology shapes taxonomic and functional structure of desert stream invertebrate communities. Freshwater Science 34: 399-409. 10.1086/680518.

Sengupta A., E.D. Stein, K. McCune, R.D. Mazor, B.P. Bledsoe, S. Adams, C.P. Konrad. (In review). From Gaged to Ungaged: Predicting Flow and Hydrologic Alterations at a Regional Scale for Application to Regional Flow-ecology analyses. Freshwater Biology

Sivapalan, M., K. Takeuchi, S.W. Franks, V.K. Gupta, H. Karambiri, V. Lakshmi, X. Liang, J.J. McDonnell, E.M. Mendiondo, P.E. O'connell and T. Oki. 2003. IAHS Decade on Predictions in Ungauged Basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences.*Hydrological sciences journal*, *48*(6), pp.857-880.

Sleeter B, T. Wilson, C. Soulard and J. Liu. 2011. Estimation of late twentieth century land-cover change in California. Environmental Monitoring and Assessment 173: 251-266. 10.1007/s10661-010-1385-8.

Solans, M.A. and D.G. Jalon. 2016. Basic tools for setting environmental flows at the regional scale: application of the ELOHA framework in a Mediterranean river basin. Ecohydrology. DOI: 10.1002/eco.1745

Souza F., F. Assis, and U. Lall. 2003. Seasonal to interannual ensemble streamflow forecasts for Ceara, Brazil: Applications of a multivariate, semiparametric algorithm. *Water Resources Research* 39, no. 11

Stein, E.D., A. Sengupta, R.D. Mazor, K. McCune, B.P. Bledsoe, and S. Adams. In review. 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.

Stein, E.D., A. Sengupta, R. Mazor, and K. McCune. 2016. Application of Regional Flow-ecology to Inform Management Decision in the San Diego River Watershed. Southern California Coastal Water Research Project Technical Report #948.

Stevens, D.L. and A.R. Olsen. 2004. Spatially balanced sampling of natural resources. Journal of the American Statistical Association 99:262-278.

Stewardson M.J. & Gippel C.J. 2003 Incorporating flow variability into environmental flow regimes using the flow events method. *River Research and Applications*, **19**(5-6), 459-472.

Stoddard, J.L., D.P. Larsen, C.P. Hawkins, R.K. Johnson and R.H. Norris. 2006. Setting expectations for the ecological condition of streams: the concept of reference condition. *Ecological Applications*, *16*(4), pp.1267-1276.

Stohlgren, T.J., T.N. Chase, R.A. Pielke, T.G.F. Kittel, and J. Baron. "Evidence that local land use practices influence regional climate, vegetation, and stream flow patterns in adjacent natural areas." *Global Change Biology* 4, no. 5 (1998): 495-504.

Strobl, C., T. Hothorn and A. Zeileis. 2009. Party on!: a new, conditional variable-importance measure for random forests available in the party package. The R Journal 1: 14-17.

U.S. Army Corps of Engineers (ACOE). 2000. Hydrologic Modeling System HEC-HMS U.S. Army Corps of Engineers R&D Technical Reference Manual.

U.S. Army Corps of Engineers (USACE), The Nature Conservancy, and the Interstate Commission on the Potomac River Basin. 2013. Middle Potomac River Watershed Assessment:

U.S. Environmental Protection Agency. 1990. Biological Criteria: National program guidance for state managers. EPA 440-4-90-010. Office of Water. Washington, D.C.

Van Sickle, J., & Hughes, R. M. (2000). Classification strengths of ecoregions, catchments, and geographic clusters for aquatic vertebrates in Oregon. *Journal of the North American Benthological Society*, *19*(3), 370-384.

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

Wagener, T. and H.S. Wheater. 2006. Parameter estimation and regionalization for continuous rainfall-runoff models including uncertainty. *Journal of Hydrology*, *320*(1), pp.132-154.

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

Walters, A.W. 2011. Resistance of aquatic insects to a low-flow disturbance: Exploring a trait-based approach. Journal of the North American Benthological Society 30: 346-356.

Wang, L., J. Lyons, P. Kanehl and R. Bannerman. 2001. Impacts of urbanization on stream habitat and fish across multiple spatial scales. *Environmental Management*, *28*(2), pp.255-266.

Webb, J.A., N.R. Bond, S.R. Wealands, R. Mac Nally, G.P. Quinn, P.A. Vesk and M.R. Grace. 2007. Bayesian clustering with AutoClass explicitly recognises uncertainties in landscape classification. Ecography 30: 526-536. DOI 10.1111/j.2007.0906-7590.05002.x.

Wickham, H. 2009. Elegant graphics for data analysis. Springer-Verlag. New York.

Wigington, P.J., S.G. Leibowitz, R.L. Comeleo, and J.L. Ebersole. 2013. Oregon Hydrologic Landscapes: A Classification Framework1. JAWRA Journal of the American Water Resources Association 49: 163-182. 10.1111/jawr.12009.

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-Journal Des Sciences Hydrologiques 59: 605-617. Doi 10.1080/02626667.2014.880786.

Wolock, D.M., T.C. Winter and G. McMahon G. 2004. Delineation and evaluation of hydrologic-landscape regions in the United States using geographic information system tools and multivariate statistical analyses. Environmental Management, Supplement 34: S71-S88. DOI 10.1007/s00267-003-5077-9.

Yadav, M., T. Wagener and H. Gupta. 2007. Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins. *Advances in Water Resources*, *30*(8), pp.1756-1774.

Yan, B., N.F. Fang, P.C. Zhang and Z.H. Shi. 2013. Impacts of land use change on watershed streamflow and sediment yield: an assessment using hydrologic modelling and partial least squares regression. *Journal of Hydrology* 484: 26-37.

Yarnell, S.M., G.E. Petts, J.C. Schmidt, A.A. Whipple, E.E. Beller, C.N. Dahm, P. Goodwin, J.H. Viers. 2015. Functional Flows in Modified Riverscapes: Hydrographs, Habitats and Opportunities. BioScience: doi: 10.1093/biosci/biv102