THESIS

ECOLOGICALLY-FOCUSED CALIBRATION OF HYDROLOGICAL MODELS FOR ENVIRONMENTAL FLOW APPLICATIONS

Submitted by

Stephen K. Adams

Department of Civil and Environmental Engineering

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Fall 2015

Master's Committee:

Advisor: Brian P. Bledose

N. LeRoy Poff
Eric D. Stein

ABSTRACT

ECOLOGICALLY-FOCUSED CALIBRATION OF HYDROLOGICAL MODELS FOR ENVIRONMENTAL FLOW APPLICATIONS

Hydrologic alteration resulting from watershed urbanization is a common cause of aquatic ecosystem degradation. Developing environmental flow criteria for managing the effects of urbanization and other human influences requires quantitative flow-ecology relationships that link biological responses to streamflow alteration. To the extent possible, gaged flow data are used; however, bioassessment sites are frequently ungaged and hydrological models must be used to characterize flow alteration. Physically-based rainfall-runoff models typically utilize a "best overall fit" calibration criterion, such as the Nash-Sutcliffe Efficiency (NSE), that does not focus on specific aspects of the flow regime relevant to biotic endpoints. This study aims to identify how accurately coastal southern California rainfall-runoff models can be calibrated using specific elements of the flow regime known *a priori* to be critical to benthic macroinvertebrates (ecologically-focused) versus a traditional best overall fit criterion. Additionally, this study seeks to assess the utility of ecologically-focused calibrated models by comparing flow metric accuracy and the strength of flow-ecology relationships among different calibration approaches versus gage data.

For this study, continuous HEC-HMS 4.0 models were created for 19 coastal southern California watersheds and calibrated to USGS streamflow gages with nearby bioassessment sites using one best overall fit and three ecologically-focused criteria: NSE, Richards-Baker Flashiness Index (RBI), percent of time when the flow is < 28 L/s (< 1 cfs), and a Combined

Calibration (RBI and < 1 cfs), respectively. Ecologically-focused criteria were selected based on preliminary statistical flow-ecology relationships at gaged bioassessment sites. Calibrated models were compared using flow metric accuracy relative to gage data and the strength of flow-ecology relationships. Models were highly accurately calibrated to ecologically-focused criteria, with calibration median percent errors less than 1.5% and only a single model with a percent error greater than 10%, and NSE criteria, with a median value of 0.634. Regardless of high calibration accuracy for ecologically-focused models, additional flow metrics not explicitly calibrated, especially those describing magnitude or rise and fall rates at aggregated daily time scales, were not consistently reproduced by models. Despite inaccuracies across a full suite of 71 flow metrics, low flow and flashiness metrics relevant to biotic endpoints were modeled accurately (< 20% error) and often provided stronger flow-ecology relationships than best overall fit criteria in terms of adjusted R^2 in multiple regression analyses and variance explained in random forest modeling. This was especially true when two ecologically-focused criteria were combined, suggesting the importance of multiple calibration criteria. Flow metrics from the Combined Calibration provided the strongest flow-ecology models in correlation and regression analyses compared to the other three calibration approaches, and perform similarly in random forest models. This study demonstrates that if ecologically relevant flow metrics can be identified using published literature or preliminary statistical analyses of gaged bioassessment sites prior to developing a hydrologic foundation, they can be incorporated as calibration criteria and provide stronger modeled flow-ecology relationships than exclusive use of a best overall fit criterion.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my adviser, Dr. Brian Bledsoe, for his continuous support of my graduate education. His patience, understanding, motivation, and immense knowledge have helped guide and improve my experience. I would also like to thank the other members of my committee, Dr. Eric Stein and Dr. Leroy Poff, for the assistance they have provided in all aspects of this research.

At the University of Florida, I would like to acknowledge the Environmental Engineering Sciences program, from which I received my B.S. degree, and the Fisheries and Aquatic Sciences Program, which first employed me as a researcher. Specifically, I'd like to thank Dr. Edward Phlips, Dr. Michael Annable, Dr. Timothy Townsend, and Dr. James Jawitz for introducing me to academic research and facilitating my success.

I would also like to thank my family and friends for the support they have provided me throughout my life, and especially during my time in graduate school. In particular, I would like to recognize my girlfriend and best friend, Kathryn. Without her love, encouragement, and sacrifice I would not have finished this thesis.

Finally, I recognize that this research would not have been possible without financial assistance from The Southern California Coastal Water Research Project and Colorado State University, for which I am extremely grateful.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

Symbols:

Q	=	flow (ft^3/s)
Q_o	=	observed flow (ft^3/s)
Q_m	=	modeled flow (ft^3/s)
$\bar{Q}$	=	mean flow (ft^3/s)
t	=	time step (hr or day)
T	=	final time step (hr or day)
TOC	=	time of concentration (min)
K	=	unit conversion (0.0078 for US units and 0.0195 for SI units)
L	=	channel flow length (ft, m, or mi)
S	=	channel slope (m/m)
R	=	Clark Unit Hydrograph storage coefficient (hr)
A	=	basin area (mi^2 or km^2)
N	=	recommended minimum number of precipitation gages per model

Statistical Terms:

n	=	sample size
p	=	p value
R^2	=	coefficient of determination

UNITS OF MEASURE

cfs	cubic feet per second
L/s	liter(s) per second
cm	centimeter(s)
m	meter(s)
m/m	meter(s) per meter
km	kilometer(s)
km^2	square kilometer(s)
in	inch(es)
in/hr	inch(es) per hour
ft	foot or feet
ft/ft	foot or feet per foot or feet
mi	mile(s)
mi^2	square mile(s)
s	second(s)
min	minute(s)
hr	hour(s)
yr	year(s)

CHAPTER 1: INTRODUCTION

Stream biota are fundamentally influenced by flow variability, a key control and indicator of ecosystem health (Poff *et al.*, 1997; Bunn and Arthington, 2002). Watershed land uses such as urbanization, agriculture, dams, and diversions have significantly modified flow regimes in streams around the world (Poff *et al.*, 1997; Walsh *et al.*, 2005; Konrad and Booth, 2005; Poff *et al.*, 2006a; Poff *et al.*, 2007). These human influences produce a wide variety of ecosystem responses, including water quality degradation, habitat loss, and increases of invasive species, but many can be traced back to flow alteration as an important causal factor (Jacobson *et al*., 2001; Bunn and Arthington, 2002; Konrad and Booth, 2005; Poff *et al*., 2006a). This understanding has led to the development of environmental flow criteria frameworks, such as the Ecological Limits of Hydrologic Alteration (ELOHA; Poff *et al.*, 2010), in which streamflow management objectives can be developed to support conservation of aquatic biota in the midst of regional hydrologic alteration.

Coastal southern California provides an excellent opportunity for the development of environmental flow criteria using the ELOHA framework given the extent of land use changes and subsequent streamflow alterations that have occurred. Hydromodification, defined as changes in channel form associated with streamflow and sediment alterations due to land use change (Stein *et al.*, 2012), caused by urbanization has altered the flow and sediment regimes of streams throughout coastal southern California, prompting the recommendation of the ELOHA framework as a method for mitigating regional hydrologic alteration (Stein *et al.*, 2012). Urbanization will further expand as California anticipates a 33% increase in population by 2060 to over 51 million residents (CA Dept. of Finance, 2014) with urban growth outpacing the overall US average (U.S. Census Bureau, 2012). Rapid urbanization has been shown to significantly increase the magnitude and duration of flows in coastal southern California streams (Hawley and Bledsoe, 2011). Land use change has also transformed California into the largest agricultural state in country (USDA, 2015). Streams in the region vary from small, ephemeral arroyos to large, fully-perennial rivers, and experience diverse ranges of urbanization influence. Despite the extent of the human footprint in coastal southern California, there hasn't been a formal effort to link flow alteration and the biological status of streams within an environmental streamflow framework. Previous studies of identifying significant associations between invertebrates and flow characteristics bode well for such an effort (Gasith and Resh, 1999; Eberhart, 2014).

Establishing a hydrologic foundation and flow-ecology relationships between streamflow departures and biological indicators such as benthic macroinvertebrates, fishes, algae, or riparian vegetation is fundamental to the development of environmental flow criteria in the ELOHA framework (Poff *et al.*, 2010). A regional hydrologic foundation requires the generation of pre- and post-alteration streamflow statistics (hereafter referred to as flow metrics), such as mean October flow or number of days per year with no flow, over some period at sites with biological data in order to assess the influence of flow departures on stream biota. Preliminary flow-ecology hypotheses are often developed from literature and expert opinion prior to statistical testing using streamflow data at gaged bioassessment sites (Konrad *et al.*, 2008; Poff *et al.*, 2010; Kendy *et al.*, 2012; Kennen *et al.*, 2013). The ELOHA framework suggests developing testable flow-ecology hypotheses independent of the hydrologic foundation using descriptive streamflow metrics that likely have a mechanistic relationship with biota and are amenable to management (Poff *et al.*, 2010).

Within the context of environmental flow criteria development, hydrologic modeling, such as rainfall-runoff modeling, is often used to predict streamflow at bioassessment sites where no observable streamflow data exist (Poff *et al.*, 2010). In the US, streamflow gages are sparse and biased toward larger rivers with 95% of streams containing less than 3% of streamflow gages. Furthermore, greater than 93% of US stream lengths are described by less than 1/3 of streamflow gages (Poff *et al.*, 2006a). Physically-based models can be used to predict flow at any point in a stream network based on precipitation data, watershed characteristics, and calibration to observed streamflow data. Furthermore, continuous simulations using rainfall-runoff models have the advantage of producing time-series of discharge data at many temporal resolutions, from which any flow characteristic can be described statistically, given a long enough time-series. These continuous, physically-based rainfall-runoff models are often more robust in characterizing human influences such as urbanization because they directly represent changes in infiltration and other hydrologic processes, as opposed to using coarse surrogates such as "percent urban land use" that are typically used in statistical or regression models. This aspect also makes physically-based models more flexible in depicting management scenarios that focus on changing specific hydrologic processes, as opposed to gross categories of land use.

Calibration of hydrologic models is a common approach in which data limitations prevent direct calculation of some parameters, and so they are adjusted to improve accuracy according to observed streamflow data. For a hydrologic foundation, accurately calibrated models can help ensure fidelity to highly variable streamflow behavior, otherwise causal linkages between flow and biological response may be missed. Rainfall-runoff model calibration is almost always performed using "best overall fit" performance measures, such as the Nash-Sutcliffe Efficiency (NSE; Nash and Sutcliffe, 1970) or error variance, despite their known bias towards large flows, disregard for timing, and residual autocorrelation (Beven, 2012; Blosch *et al.*, 2013), and inability to capture flow metrics of biological interest, such as the seven-day minimum flow and timing of annual minimum runoff (Cassin *et al.*, 2005; Vis *et al.*, 2015). Rainfall-runoff models calibrated to only one criterion have been shown to be less likely to accurately predict a range of ecological flow characteristics (Murphy *et al*., 2013; Vis *et al.*, 2015). Beven and Binley (1992) argue that rainfall-runoff models should be created for only specific purposes, which should be known *a priori* and kept in mind when applying the model outputs. Likewise, it has been suggested that prior selection of ecologically relevant streamflow characteristics and targeted calibration can improve the application of hydrologic models in ecological flow studies (Cassin *et al.*, 2005; Murphy *et al*., 2013); however, this has not been tested to my knowledge. For the purpose of developing environmental flow criteria, a more robust approach to hydrologic model calibration is needed—one focused on elements of the flow regime known to be important to biological endpoints using prior knowledge of flow-ecology relationships—such that a modeled hydrologic foundation is more relevant to manageable biotic endpoints.

The creation of an accurate hydrologic foundation is crucial for establishing environmental flow criteria and flow-ecology relationships in the wake of urbanization in coastal southern California, a method that has been recommended for combatting hydromodification in the region (Stein *et al.*, 2012). These issues pose the primary research problem of my study: How can a regional set of physically-based rainfall-runoff models be created and calibrated to resolve critical elements of human-induced hydrologic alteration across gradients of flow intermittency and urbanization using prior knowledge of regional flow-ecology relationships for benthic macroinvertebrates for application as a hydrologic foundation? In this study, the feasibility of utilizing an “ecologically-focused” calibration approach at a sub-daily time scale in the development of a regional hydrologic foundation is explored. The specific objectives are to:

1) Identify how accurately coastal southern California rainfall-runoff models can be calibrated using specific elements of the flow regime critical to benthic macroinvertebrates (ecologically-focused) versus traditional best overall fit criteria;
2) Explore how best overall fit and ecologically-focused calibration methods affect the accuracy of rainfall-runoff models that will be used for a hydrologic foundation; and
3) Compare how well hydrologic metrics derived from calibrated rainfall-runoff models explain variability in benthic macroinvertebrate assemblages versus metrics calculated directly from streamflow records at gaged sites.

In addressing these objectives, I will develop and describe a novel approach to creating hydrologic models that more accurately simulate key elements of the flow regime known to strongly influence regional benthic macroinvertebrate assemblages as compared to models calibrated based solely on overall fit criteria. To achieve the study objectives, flow characteristics known *a priori* to significantly shape biological assemblages in coastal Southern California (Eberhart, 2014) and the NSE were used as competing calibration criteria for 19 locally-calibrated rainfall-runoff models with outlets at U.S. Geological Survey (USGS) streamflow gages matched to nearby bioassessment sites in coastal southern California. Flow metrics are computed from both gage data and rainfall-runoff models subjected to the differing calibration methods. Model results are then subjected to a statistical assessment of accuracy, and ultimately confronted with biological metrics calculated from nearby benthic macroinvertebrate bioassessment data to assess their relative explanatory power.

CHAPTER 2: BACKGROUND

2.1 Natural and altered flow regimes

The ecological structure and function of streams is fundamentally controlled by multiple aspects of the flow regime including flow magnitude, frequency, duration, timing, and rate of change (Poff and Ward, 1989). These flow regime components directly influence numerous factors critical to ecological integrity such as water chemistry, geomorphic processes, habitat structure, and biological organization (Poff *et al*., 1997; Bunn and Arthington, 2002) making them the focus of many environmental flow frameworks (Poff *et al.*, 2010). Natural flow regimes vary tremendously across US regions, ranging from perennial flows driven by seasonal rainfall variability in the humid southeast, to flashy ephemeral streams in the arid west, to relatively stable and predictable snowmelt-dominated systems at high elevations (Poff, 1996). Each of these regions contains biota adapted to the local flow regime. Based on this understanding, regional flow-ecology relationships and environmental flow criteria can be developed to incorporate aspects of the flow regime most critical to local flora and fauna (Poff *et al.*, 2010).

The 20th and 21st centuries have been characterized by rapid land use and water resource infrastructure changes that have significantly altered natural flow regimes. Urbanization growth has been shown to increase the magnitude and frequency of large flows, which can incise, widen and enlarge channels (Booth, 1990; Bledsoe and Watson, 2001; Hawley *et al*., 2012). In attempts to prevent flooding and undesirable geomorphic adjustments in agricultural or urban areas, streams are often fixed in place with levees and channel armoring, which can lead to channel homogeneity, habitat loss and reduced interaction with riparian ecosystems. Presently, urbanization is perhaps the most ubiquitous driver of hydrologic alteration due to its rapid increase at a global scale (Forman, 2008). Flashiness and low flow magnitudes and frequencies are known to be susceptible to urbanization as streams become flashier and baseflows may increase or decrease, often depending on the prevalence of effluent discharge and other factors (Walsh *et al*., 2005). In a southern California case study, the peak and duration of all sediment-mobilizing flows increased in response to urbanization such that a typical watershed with 20% impervious area could increase the two year recurring flow six fold compared to areas with no imperviousness (Hawley and Bledsoe, 2011). Indeed, the direct and indirect effects of urbanization and other land use changes have been shown to affect all five elements of the natural flow regime (magnitude, frequency, duration, timing, and rate of change; Poff *et al.*, 1997). This poses a substantial problem for the many scientists, engineers, and managers working towards creating safe, functional urban water systems while simultaneously maintaining ecological integrity. The recent attention to environmental flow standards has generated a rapid increase in the numbers of studies focused on improving the ecological integrity of streams and watersheds through flow management (TNC, 2015).

Quantitative flow metrics describing regional hydrologic variability are often calculated from regional regression equations or time-series of streamflow data to compactly and statistically describe various elements of the flow regime. Time-series of streamflow data from gages or rainfall-runoff models are advantageous because they contain raw data from which numerous metrics may be calculated, while regional regression equations are calibrated to calculate only one specific metric. Software such as Indicators of Hydrologic Alteration (IHA; Richter *et al*., 1996), GeoTools (Bledsoe *et al*., 2007), and the Hydroecological Integrity Assessment Process' National Hydrologic Assessment Tool (NATHAT; Henriksen *et al*., 2006) are frequently used to compute descriptive flow metrics.

Additional metrics not computed with software packages that describe streamflow flashiness, such a T_{Qmean} (Konrad *et al.*, 2005) and Richards-Baker Flashiness Index (RBI; Baker *et al*., 2004), are used in flow-ecology studies due to their known biological relevance in systems with variable streamflow regimes. The RBI is widely used (Baker *et al*., 2004):

$$RBI = \frac{\sum_{t=1}^{T} |Q_{t+1} - Q_t|}{\sum_{t=1}^{T} Q_t} \qquad \textbf{Eqn. 2.1}$$

wherein Q_t is the discharge at time t, Q_{t+1} is the discharge at time step after t, and T is the final time step. T_{Qmean} quantifies the fraction of time that flow exceeds the mean streamflow for a specific duration. RBI increases with flashiness; in contrast, the value of T_{Qmean} decreases as flashiness increases (Konrad *et al*., 2005). RBI is dependent on the temporal density of flows, so only values calculated at the same time step (e.g. hourly) can be compared for relative flashiness (Baker *et al*., 2004).

When time-series of streamflow data are used to compute flow metrics, traditionally resolutions of daily, weekly or even monthly time steps are used; however; such coarse temporal densities have proven ineffective in capturing flow-ecology relationships in North Carolina (Pomeroy *et al.*, 2008) and in describing urbanization effects in North Carolina and Wisconsin (Knight and Cuffney, 2012), where 15-minute and hourly time steps, respectively, were more effective. The importance of data resolution for capturing streamflow alterations within hydropower ecohydrological studies has been specifically investigated recently and it was found that flow metrics taken from sub-daily data explain variation in hydrology among streams and provide more noticeable variation within streams than daily data for natural "run-of-river" and altered hydropower "peaking" flow regimes (Bevelhimer *et al.*, 2014).

2.2 Benthic macroinvertebrates in environmental streamflow applications

Benthic macroinvertebrates are the most commonly used bioassessment endpoint for stream and watershed management and monitoring programs (Resh *et al.*, 2006). Dating back to the early 20th century (Kolwitz and Marsson, 1909), they have been used as key indicators of water chemistry and aquatic habitat quality (Cairns and Pratt, 1993). The US EPA's Rapid Bioassessment Protocol indicates most state water quality agencies that regularly survey biological data focus on benthic macroinvertebrates and recommends their use in monitoring because invertebrates represent highly localized site-specific regions of streams due to their immobile nature relative to other creatures such as birds and fish (Barbour *et al.*, 1999). This is especially useful for establishing flow-ecology relationships when streamflow data are limited to sparsely distributed gage sites or specifically modeled locations in a stream network. Benthic macroinvertebrates serve critical ecosystem functions and comprise the foundation of food webs for many important fish and bird species. Collection and analysis of benthic macroinvertebrates is relatively simple, quick, and inexpensive compared to other measures of water quality, such as suites of chemical analyses (Barbour *et al.*, 1999; Resh *et al.*, 2006). Benthic macroinvertebrates are also appropriate for developing flow-ecology relationships and environmental flow criteria because they detectably respond to hydrologic alteration created by land use change (Konrad *et al.*, 2008; Poff and Zimmerman, 2010; Kennen and Riskin, 2010; Kennen *et al.*, 2010; Brooks *et al.*, 2011; Kennen *et al.*, 2013; Eberhart, 2014).

Previous flow-ecology studies utilizing benthic macroinvertebrates have shown increased stream flashiness associated with reduced biotic integrity in lowland streams of the US Pacific Northwest (Booth *et al*., 2004; Cassin *et al.*, 2005) and reduced Ephemeroptera, Plecoptera, and Trichoptera (EPT) richness in Piedmont streams of the southeast US (Pomeroy *et al.*, 2008).

Studies on the response of benthic macroinvertebrates to seasonal flows in Mediterranean regions, including coastal southern California and Europe, have revealed that assemblages are critically shaped by minimum flows and predictable drying events (Gasith and Resh, 1999; Datry, 2012; Belmar *et al.*, 2013). Streamflow time-series of sub-daily time resolution have been shown to more accurately describe benthic macroinvertebrate assemblages (Helms *et al.*, 2009), often specifically citing the relationship between stream flashiness and benthic macroinvertebrates (Cassin *et al.*, 2005; Pomeroy *et al.*, 2008; Bevelhimer *et al.*, 2014); however, they have seldom been used in developing hydrologic foundations for environmental flow criteria studies (Cassin *et al.*, 2005; Pomeroy *et al.*, 2008).

In a study investigating the appropriate time period of a hydrologic foundation for ecohydrological studies, Kennard *et al.* (2010) recommend a duration long enough to sufficiently capture the hydrologic variability that shapes stream biotic assemblages, normally around 15 years of daily streamflow data. For benthic macroinvertebrates specifically, flow metrics generated from both long-term (5 to 15 years) and short-term (30 to 100 days) periods prior to sampling have been shown to explain significant variation in invertebrate assemblages (Konrad *et al.*, 2008). Historical ecohydrological studies involving benthic macroinvertebrates have focused on periods of streamflow data less than three years (Likens, 1984), with only 12% of 266 selected research articles between 1980 and 1987 using durations greater than three years (McElravy, 1988). Other, more recent studies have found that a three year hydrologic record is adequate to describe regional benthic macroinvertebrate variability (Kennen *et al.*, 2010; Eberhart, 2014).

2.3 Coastal southern California flow-ecology

Coastal southern California has a Mediterranean climate where streams are generally flashy, seasonal, and rapidly urbanizing (Gasith and Resh, 1999; Hawley and Bledsoe, 2011). Previous research indicates that stream flashiness, drying, and the duration of extreme low flows strongly influence benthic macroinvertebrates in coastal southern California (Gasith and Resh, 1999; Zickovich and Bohonak, 2007; Eberhart, 2014). This study benefits from preliminary flow-ecology relationships established at USGS streamflow gages using nearby benthic macroinvertebrate bioassessment sites in the region (Eberhart, 2014). Statistical analyses indicate that the frequency of flows greater than 28 L/s (~1 cfs) and flashiness explain the most variability in a set of taxonomic and trait-based benthic macroinvertebrate metrics. The 28 L/s (~1 cfs) threshold was used in the preliminary study as a proxy for stream drying due to the sensitivity and uncertainty associated with measuring very small flows at USGS gages. In particular, percent richness of non-insect taxa, EPT taxa, shredders, taxa resilient to disturbance, taxa resistant to desiccation, and taxa resistant to sand-bed instability were significantly correlated ($p = 0.1$) with the percent of time above a low flow threshold. Flow flashiness was a significant predictor of a multi-metric index of biotic integrity specifically created for southern California (SC-IBI; Ode *et al.*, 2005), shredder richness resilience to disturbance, and resistance to sand-bed instability traits (Eberhart, 2014). In ELOHA-based and other environmental flow studies, it is fairly common to develop such preliminary flow-ecology relationships prior to the hydrologic foundation, especially through hypotheses; however, I could not identify any scientific literature documenting the use of *a priori* biological knowledge to guide hydrological model calibration for development of a hydrologic foundation beyond suggestions that it should be explored (Cassin *et al.*, 2005; Murphy *et al*., 2013). These principles highlight the primary rationale for this study.

Many ecohydrological studies utilize long records of streamflow, often around 15 years, to develop regional hydrologic foundations (Kennard *et al.*, 2010); however, perhaps due to their relatively short life-cycles compared to other biota used to develop environmental flow criteria, such as fishes, some studies have concluded through sensitivity analyses that a three year streamflow record balances sufficient representation of temporal hydrologic variability that shape benthic macroinvertebrate assemblages with data availability for a sufficient number of (Kennen *et al.*, 2010; Eberhart, 2014). As aforementioned, sub-daily time steps are ideal for capturing the effects of urbanization and stream flashiness in coastal southern California and other flashy hydrologic settings; however, the tremendous data management and processing hurdle of using sub-daily data is demonstrated by some simple arithmetic: 15 years of daily data (15*365=5,475) contains nearly 5 times *fewer* data points than 3 years of hourly data (3*24*365=26,280). In order to accurately predict sub-daily streamflow using a rainfall-runoff model, sub-daily precipitation input data must be used. This again poses significant processing and management challenges, as most sub-daily precipitation data are not correctly formatted for rainfall-runoff models and must be carefully scrutinized for commonplace errors. These issues make it very laborious to incorporate sub-daily data into long-term regional hydrologic foundations, and are perhaps why sub-daily modeling has not been frequently pursued within the ELOHA framework. For these reasons, it seems appropriate to utilize a high resolution (hourly) hydrologic foundation, at the expense of a shorter than traditional duration (three yr) that has still produced flow metrics shown to be significant predictors of benthic macroinvertebrate metrics in coastal southern California and other regions for environment flow criteria development.

2.4 Rainfall-runoff modeling

Physically-based rainfall-runoff models simulate streamflow using rainfall data and mathematical equations representing simplified physical processes tracking water movement through a watershed. As a hydrologic foundation for flow-ecology studies, rainfall-runoff models may be used to predict time-series of streamflow and flow metrics at bioassessment sites (Poff *et al.*, 2010). The use of physically-based and readily measureable or calculable parameters in rainfall-runoff models, such as the percent of basin area covered in impervious surfaces or time of concentration, make them an attractive option for directly characterizing human influences and management actions on streamflow through time. Parameters within rainfall-runoff models can be adjusted based on physical understanding to represent different land use and stormwater management scenarios. This allows for simple computation of hydrologic alteration using unaltered- and altered-condition parameter sets, and offers more flexible management options for environmental flow application due to an increased ability to describe hydrologic alteration and compare different management scenarios (Kendy *et al*., 2012).

Some rainfall-runoff models are designed so that all parameters can be estimated from field-collected data. Manually measuring all parameters can be a very resource intensive process; thus, rainfall-runoff models are commonly calibrated using existing streamflow gage data (Beven, 2012). When performed responsibly, calibration provides a sound method for improving the performance of rainfall-runoff models. In practice, calibration involves automatically or manually altering model parameters until one or more goodness of fit criteria are satisfied. Unfortunately, calibration criteria used in flow-ecology studies are not often described in a transparent manner in the hydrologic literature. In studies that do report calibration methods, a single best overall fit criterion, such NSE (Nash and Sutcliffe, 1970), is most often used for calibration (Jain and Sudheer, 2008; Beven, 2012; Bloschl *et al.*, 2013; Vis *et al.*, 2015).

The NSE has become one of the most widely-used criteria for calibrating rainfall-runoff models, despite evidence showing high values of NSE (> 0.5) can be obtained even when the fit is relatively poor and vice-versa (Jain and Sudheer, 2008). These problems arise when the variance of streamflow values is large, leading to relatively larger residuals for high flows and resulting in bias fit emphasizing high flows. By definition, NSE assesses the accuracy of a model prediction relative to the observed mean flow during the model duration, such that a modeled time series containing only the observed mean flow repeated for every time step would have an efficiency of zero (Nash and Sutcliffe, 1970). The worst possible NSE value is negative infinity, while a perfect model that identically matches the observed discharges has an NSE of one. The formula for NSE, as defined by Nash and Sutcliffe (1970), is:

$$NSE = 1 - \frac{\sum_{t=1}^{T}\left(Q_o^t - Q_m^t\right)^2}{\sum_{t=1}^{T}\left(Q_o^t - \overline{Q_o}\right)^2} \qquad \textbf{Eqn. 2.2}$$

wherein $\overline{Q_o}$ is the mean of observed discharges, $Q_m{}^t$ is the modeled discharge at time *t*, $Q_o{}^t$ is the observed discharge at time *t*, and *T* is the final time step. By definition, NSE does not explicitly incorporate model accuracy for different elements of the flow regime. In many stream management applications model accuracy of key aspects of the flow regime, such as summer baseflow for sensitive biota or peaks for flood planning, are much more critical than overall accuracy relative to the mean flow (Beven and Binley, 1992; Cassin *et al.*, 2005). Rainfall-runoff models calibrated to only one criterion, such as NSE, are less likely to accurately predict a range of ecological flow characteristics (Murphy *et al*., 2013; Vis *et al.*, 2015), and the use of multiple calibration criteria is recommended for more robust calibration (Gupta *et al.*, 2008). The accuracy of ecologically relevant metrics vary with different best overall fit criteria (Vis *et al.*, 2015). In order to establish a useful hydrologic foundation for developing flow-ecology relationships, modeled ecological flow metrics specifically relevant to regional bioassessment endpoints need to be accurate.

For the foregoing reasons, more rigorous calibration and testing approaches are needed to generate accurate and ecologically relevant streamflow predictions in regional environmental flow studies. Prior knowledge often exists regarding the influence of key flow alterations on the ecological response of regional streams (Resh *et al.*, 1988). If flow metrics that substantially influence biological endpoints can be hypothesized or statistically demonstrated early in the process of developing environmental flow criteria, they could also be incorporated as guidelines for calibration of rainfall-runoff models that are accurate for the most flow regime elements that most affect the stream biota of interest. It is unknown how much the accuracy of rainfall-runoff models could improve when calibrated specifically for the most ecologically relevant components of the flow regime. Furthermore, I have not seen a previous study that compares the extent of variability in regional benthic macroinvertebrate assemblage data explained by calibrated rainfall-runoff models versus gage data.

CHAPTER 3: METHODS

3.1 Study area

This study utilizes 25 benthic macroinvertebrate bioassessment sites sampled under the Surface Water Ambient Monitoring Program (SWAMP) and the Southern California Stormwater Monitoring Coalition (SMC), and 19 sub-daily USGS streamflow gages in the coastal southern California region. To establish preliminary flow-ecology relationships, each of the bioassessment sites used in this study was matched to 14 of the sub-daily USGS streamflow gages such that the flow regime at each bioassessment site is represented by a nearby sub-daily streamflow gage that meets criteria detailed in Eberhart (2014). In six instances, multiple bioassessment sites on the same stream were matched to a single streamflow gage. To bolster the regional hydrologic foundation developed in this study, five streamflow gages were included without available matched bioassessment data; however, benthic macroinvertebrate samples have been collected or are planned near these five gage sites. The 25 bioassessment and 19 stream gage sites are shown in Figure 3.1.

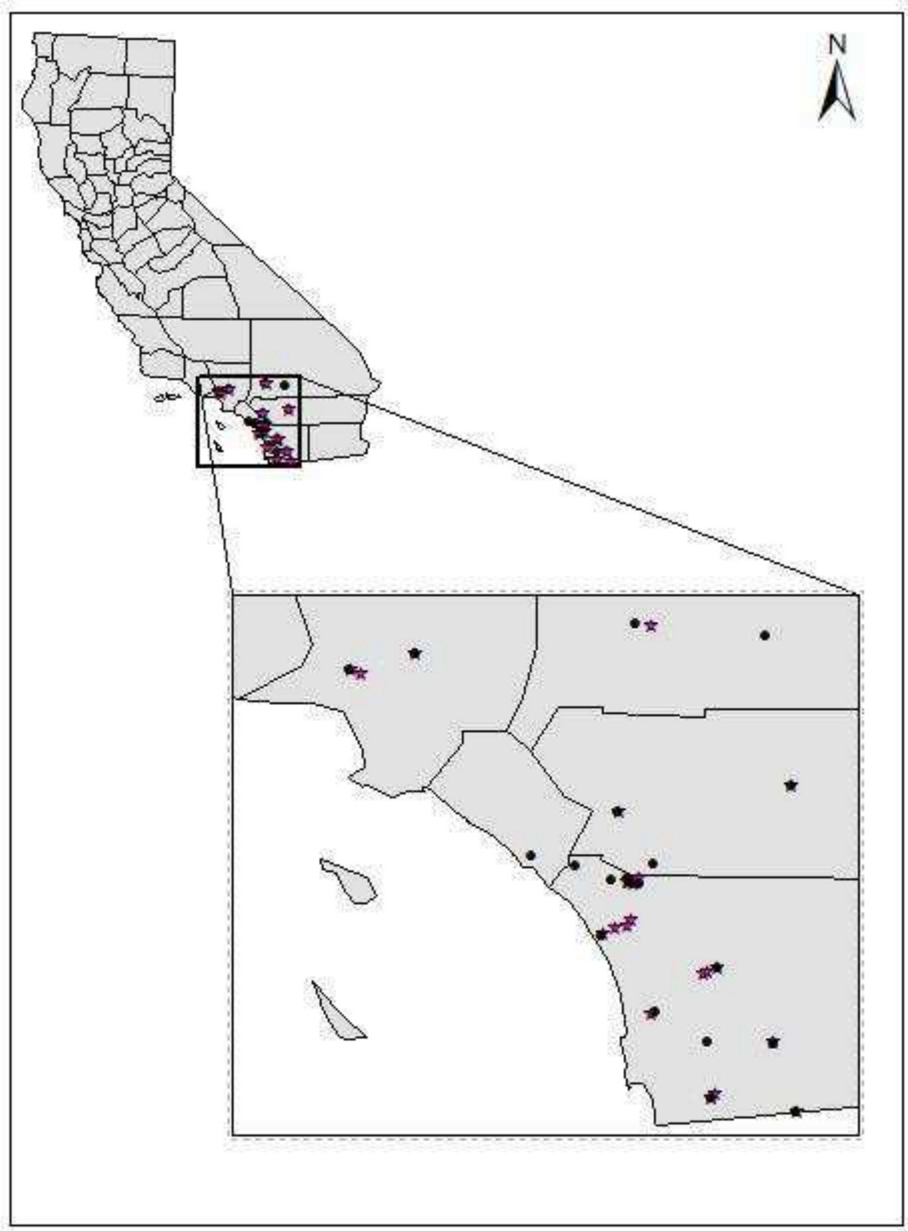

Figure 3.1: Coastal southern California study area in which circles represent streamflow gage location and each star indicates a bioassessment site.

The study area encompasses coastal regions of San Diego, Riverside, Orange, San Bernardino, and Los Angeles Counties, with some basin boundaries extending into Ventura County. The area is bounded by the Transverse Ranges to the north, Mexico to the south, the Peninsular Ranges to the east, and the Pacific Ocean to the west. Sites are characterized as the “south coast” California hydrologic region (Gotvald *et al.*, 2012), and fall within the southern and Baja California pine-oak mountains, and California coastal sage, chaparral, and oak woodlands Level III ecoregions (EPA, 2014).

The climate of the study area is Mediterranean with hot, dry summers and mild, wet winters. Wildfires are fairly common in the summer. Stream types include perennial, intermittent, and ephemeral with both sand-dominated and gravel/cobble-dominated substrates. The study region spans a broad gradient of urbanization intensity with percent impervious surfaces within streamflow gage contributing areas ranging from 0 - 27% (Fry *et al.*, 2011). Streams vary from minimally impacted by contemporary human land uses to engineered stormwater channels that are fully encased concrete.

In addition to the matching criteria provided by Eberhart (2014), study sites were chosen to represent the heterogeneity of coastal southern California. The contributing watershed area for each site ranges from 23.2 km^2 to 1718 km^2 with streamflow gage elevations ranging from 6.1 m to 930 m. Average basin slopes range from 14 - 57% while average annual precipitation ranges from 39 - 78 cm (Falcone, 2011). Additional basin characteristics for the 19 USGS streamflow gages are provided (Table 3.1).

Table 3.1: Range of basin characteristics for the 19 USGS gages.

Characteristic	Minimum (USGS Gage)	Maximum (Gage)	Source
Basin Area (km^2)	**22.4** (10259000 Andreas)	**1873** (11070500 San Jacinto)	Falcone, 2011
Mean Basin Elevation (m)	**262** (11023340 Poway)	**1787** (10260500 Deep Creek)	Falcone, 2011
Average Basin Slope (%)	**14** (11092450 Los Angeles)	**57** (11098000 Arroyo Seco)	Falcone, 2011
Average Annual Precipitation (cm)	**39.0** (11023340 Poway)	**78.4** (11098000 Arroyo Seco)	Falcone, 2011
Average Sand Content (%)	**33** (11047300 Arroyo Trabuco)	**64** (10260500 Deep Creek)	Falcone, 2011
Average Silt Content (%)	**24** (10260500 Deep Creek)	**48** (11098000 Arroyo Seco)	Falcone, 2011
Average Clay Content (%)	**10** (10259000 Andreas)	**25** (11047300 Arroyo Trabuco)	Falcone, 2011
Imperviousness (%)	**0** (10259000 Andreas)	**27** (11092450 Los Angeles)	Fry *et al.*, 2011
Time with no flow (%)	**0** (Ten gages)	**68** (11014000 Jamul)	USGS, 2014; USGS, 2015

3.2 Creation of rainfall-runoff models

For the 19 sub-daily USGS streamflow gage locations (Figure 3.1), rainfall-runoff models were created using the newest version of the Hydrologic Modeling System developed by the US Army Corps of Engineer's Hydrologic Engineering Center (HEC-HMS Version 4.0; Hydrologic Engineering Center, 2013). Creation of the HEC-HMS rainfall-runoff models is the first step in establishing the regional hydrologic foundation. These rainfall-runoff models produce sub-daily time series of discharge specifically at gage locations, which provide observable streamflow data to use for model calibration. A unique HEC-HMS model was created for each of the 19 gaged locations such that each model contains only the drainage area of each gage and simulates flow only at the streamflow gage located at the outlet. HEC-HMS is commonly used by practitioners in modeling and designing urban best management practices and is regarded as an industry standard modeling platform. Furthermore, HEC-HMS was chosen due to its on-going development, facility in performing sub-hourly, long-term continuous simulations with relatively few parameters (Hydrologic Engineering Center, 2013), approval by FEMA (Federal Emergency Management Agency, 2013), and public availability.

3.2.1 Data resolution and duration

To resolve ecologically important streamflow flashiness and the effects of urbanization, the 19 HEC-HMS models were set up to run at an hourly time step. Some input data, including gaged streamflow and some precipitation, were input at a 15 minute time density; however, model computations were performed at hourly time steps. Models could not be accurately run at the 15 minute time scale because hourly precipitation data were the finest temporal density available at many locations. Each model outputs an hourly time series of flow, which can be aggregated to daily data.

A regional hydrologic foundation developed for ecohydrological studies should span a duration that sufficiently represents the hydrologic variability that shapes biological endpoints (Kennard *et al.* 2010). Recent studies have suggested that three years is adequate for benthic macroinvertebrates in coastal southern California (Eberhart, 2014) and New Jersey (Kennen *et al.*, 2010); however, three years is much shorter than the typical hydrologic foundation used for ecohydrologic studies (Kennard *et al.*, 2010). A three year, hourly hydrologic foundation was utilized in this study. To control for climatic variability to the extent possible, all 19 models were created for the same time period, specifically water years (WY) 2005 - 2007. This period is represented by relatively abundant and high quality streamflow and precipitation data and characterizes regional hydrologic variability by including a wet, average, and dry year (Figure 3.2).

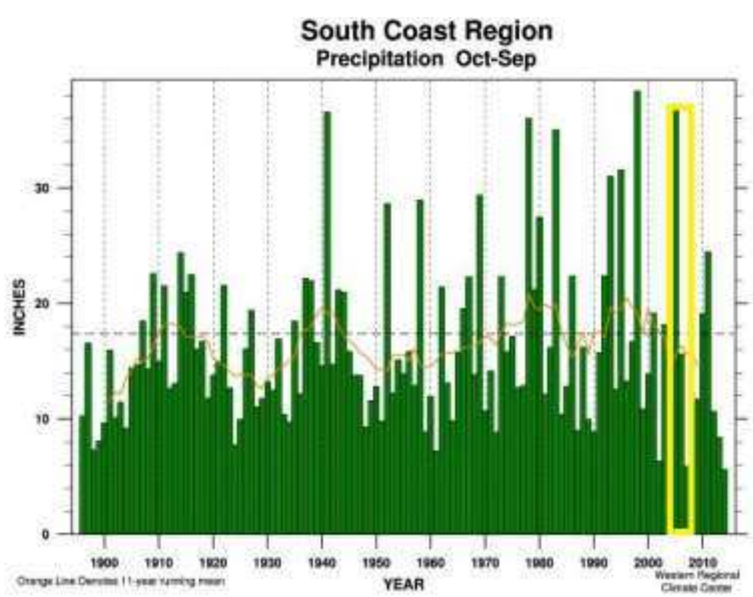

Figure 3.2: Total precipitation volume by water year for the South Coast Region of California (WRCC, 2015). WY 2005-2007 boxed in yellow to show the hydrologic variability of models is covered by a wet, average, and dry water year.

3.2.2 HEC-HMS model inputs

HEC-HMS and other rainfall-runoff simulation models utilize input parameters and mathematical relationships to convert rainfall input to streamflow output. Methods applicable to continuous models that can simulate long periods of no rainfall are needed for the three year models used in this study. To simulate infiltration losses, the simple canopy, simple surface, and deficit and constant loss methods were used. The Clark Unit Hydrograph technique was used to transform excess precipitation to surface runoff, while the linear reservoir method with two layers was used to represent baseflow contributions. Each of these methods contain a series of parameters, some of which are calculated directly from unique characteristics of each basin and others that are calibrated. Table 3.2 depicts the input methods and parameters used for the 19 models and indicates which parameters were calibrated versus calculated directly.

Table 3.2: HEC-HMS input methods and parameters wherein bold parameters were calculated directly and all others were calibrated.

Method	Parameters
Simple Canopy	• Maximum Storage (in) • Initial Storage (%)
Simple Surface	• Maximum Storage (in) • Initial Storage (%)
Deficit and Constant (Loss)	• Initial Deficit (in) • Maximum Deficit (in) • Constant Rate (in/hr) • **Imperviousness (%)**
Clark Unit Hydrograph (Transform)	• **Time of Concentration (hr)** • **Storage Coefficient (hr)**
Linear Reservoir (Baseflow)	• Ground Water (GW) 1 Initial Discharge (cfs) • GW 1 Storage Coefficient (hr) • # of GW 1 Reservoirs • GW 2 Initial Discharge (cfs) • GW 2 Storage Coefficient (in) • # of GW 2 Reservoirs

To facilitate direct estimation of model parameters based on the unique characteristics of each basin, the drainage area of each of the 19 streamflow gages was delineated in ArcGIS 10.1

using National Elevation Dataset (NED) 10 m DEM (Gesch *et al.*, 2002) and cross-checked with USGS gage annual reports. Percent imperviousness was computed for each basin by clipping 2006 NLCD data (Fry *et al.*, 2011) in ArcGIS. Time of concentration (TOC), which can be defined as the time required for water to flow from the furthest location in a watershed to the outlet, was calculated using the well-established Kirpich Method (Kirpich, 1940). The Kirpich Method equation is displayed below wherein *TOC* is the time of concentration in minutes, *K* is a unit conversion coefficient equal to 0.0078 for US units and 0.0195 for SI units, *L* is the channel flow length in feet or meters, and *S* is the channel slope over the entire watershed (m/m or ft/ft):

$$TOC = K * L^{0.770} * S^{-0.385} \qquad \textbf{Eqn. 3.1}$$

Variables used in the Kirpich method were obtained from the 10 m DEM ArcGIS delineations. The Clark Unit Hydrograph storage coefficient was calculated directly using a method agreed upon by the Colorado State Engineer (Tierra Grande International Inc., 2008) and Arizona Department of Transportation (NBS/Lowry Engineers & Planners, Inc. and George V. Sabol Consulting Engineers, Inc., 1993). The Clark Unit Hydrograph storage coefficient equation is:

$$R = 0.37 * TOC^{1.11} * L^{0.80} * A^{-0.57} \qquad \textbf{Eqn. 3.2}$$

wherein *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 in mi^2. Variables to calculate *R* were obtained from each basin delineation and the Kirpich method. The estimated parameter values for time of concentration, basin storage coefficient, percent imperviousness, and basin area are provided for the 19 modeled USGS streamflow gages (Table 3.3).

Table 3.3: Calculated model parameters and climate data information for 19 modeled USGS streamflow gages.

		Calculated Model Parameters				Climate Data		
USGS Gage	HEC-HMS Model Name	Time of Concentration (hr)	Basin Storage Coefficient (hr)	% Imperviousness	Basin Area (km^2)	CIMIS Evapotranspiration Site	Recommended Minimum # of Precipitation Gages	Number of Precipitation Gages Used
10259000	Andreas	0.92	0.52	0.00	22.40	Cathedral City #118	2	5
11098000	Arroyo Seco	1.95	1.25	0.46	41.44	Glendale #133	2	6
11047300	Arroyo Trabuco	4.10	2.30	19.06	140.17	Irvine #75	3	8
11012500	Campo	4.41	1.77	0.55	217.85	Otay Lake #147	3	3
11044800	De Luz	2.49	1.01	0.32	83.66	Temecula #62	2	9
10260500	Deep Creek	2.09	0.32	2.34	347.06	Lake Arrowhead #192 and Big Bear Lake #199	4	10
11015000	Descanso Sweetwater	3.60	1.73	0.28	117.59	Escondido SPV #153	3	8
11014000	Jamul	2.83	0.91	0.54	181.59	Otay Lake #147	3	5
11092450	Los Angeles	4.15	1.16	27.34	409.22	Chatsworth #215	4	18
11022200	Los Coches	1.54	0.69	9.39	31.52	Miramar #150 and Escondido SPV #153	2	4
11023340	Poway	2.52	0.95	20.66	109.92	Escondido SPV #153	3	5
11044250	Rainbow	1.82	1.09	3.70	26.44	Temecula #62	2	3
11070500	San Jacinto	5.05	0.59	5.86	1872.57	Winchester #179	6	13
11042000	San Luis Rey	13.45	5.62	2.97	1439.60	Temecula #62	5	23
11046300	San Mateo	3.53	1.28	0.13	209.27	Temecula #62	3	12
11044350	Sandia	2.71	1.50	1.27	50.95	Temecula #62	2	6
11044300	Santa Margarita Sump	9.40	2.72	3.76	1597.77	Temecula #62 and Winchester #179	6	19
11044000	Santa Margarita Temecula	3.11	0.27	4.11	1522.92	Temecula #62	5	15
11025500	Santa Ysabel	4.23	1.66	0.10	288.60	Escondido SPV #153	3	10

Besides precipitation, which is discussed in the next section, evapotranspiration is the final input requirement for the 19 HEC-HMS models. The California Irrigation Management Information System (CIMIS) contains a network of over 145 weather stations that record evapotranspiration. For each of the 19 gage locations, monthly average evapotranspiration values (in/month) for the CIMIS site nearest each gage location were used within each model (CIMIS, 2015). An evaporation coefficient of 0.77, calculated at nearby Lake Elsinore, was used for each model (Chow *et al.*, 1988).

3.2.3 Precipitation data

Locating, formatting, quality assuring (QA), and quality controlling (QC) sub-daily precipitation data from a variety of sources is an arduous task and is perhaps why most ELOHA studies with similar objectives have not used them. In all, the number of precipitation gages used per rainfall-runoff model range from three to twenty-three. Table 3.4 includes information about precipitation data sources and types. Occasionally, inexplicably large precipitation pulses caused by recording errors occurred for short durations throughout the precipitation data. Hence, expert judgment and cross-validation with other nearby gages were used to manually QA/QC precipitation time series.

Table 3.4: Precipitation data sources.

Source	Time-Step Resolution	Type	Number of Times Used
National Climate Data Center (NOAA NCDC)	1 hr	Incremental	47
California Irrigation and Management System (CIMIS)	1 hr	Incremental	21
California Data Exchange Center (CDEC)	1 hr	Cumulative	70
San Diego County Flood Control District (SDCFCD)	15 min	Incremental	40
Ventura County Watershed Protection District (VCWPD)	1 hr	Incremental	4

The suggested precipitation gage density requirement put forth by Schaake (1981) was used as a guideline for the minimum number of gages per model in this study:

$$N = 0.6 * A^{0.3} \qquad \textbf{Eqn. 3.3}$$

wherein N is the number of precipitation gages and A is the basin area in km^2.

Inverse distance weighting was used to weight all the precipitation gages within each model. The centroid of each basin was used as the weighting location such that precipitation gages nearest the basin centroid received the most weight in estimating basin-wide precipitation. Using inverse distance weighting allows for the inclusion of gages containing missing precipitation data. When data are missing, the gage is simply no longer considered in the weighting scheme until data resume again. This flexibility allows for the inclusion of many precipitation gages, regardless of missing data. Due to the coastal southern California climate and the lack of high elevation sites in this study, snowmelt was not simulated in the HEC-HMS models.

Climate information for the models is also provided in Table 3.3. For some basins, two CIMIS evapotranspiration sites were approximately equally distant. In these models, the two monthly CIMIS evapotranspiration data sets were arithmetically averaged. The minimum number of recommended models according to Eqn. 3.3 (Schaake, 1981) is included in Table 3.3. Because the inverse distance weighting method explicitly accounts for precipitation gage distance and missing data, which often exist in all sources, a greater number of precipitation gages than the recommended minimum were used for all models except 11012500 Campo, where the minimum was used due to a lack of any other reasonably nearby precipitation gages. Intuitively, models of larger basins typically used a greater number of precipitation gages, however sub-regional precipitation gage density played an important role in how many gages were used.

3.3 Calibration of rainfall-runoff models

Thirteen model parameters (Table 3.2), some with interdependence, were used for calibration with observed streamflow gage data. These parameters were manually adjusted to improve model accuracy due to the lack of Monte Carlo sampling within HEC-HMS and its limitation of incorporating only one objective function in automatic calibration. Previous studies suggest that the use of a single objective function supplies only enough information to automatically estimate three to five parameters, making it a poor choice for these models (Beven, 1989; Jakeman and Hornberger, 1993).

This study focuses on utilizing prior knowledge of ecologically relevant flow metrics as rainfall-runoff model calibration criteria. In addition to employing the common practice of calibrating rainfall-runoff models solely to best overall fit criteria, known ecologically relevant calibration criteria based on the initial analysis performed by Eberhart (2014) were used. As mentioned previously, Eberhart (2014) found streamflow flashiness and the percent of time when flow is greater than 28 L/s (~1 cfs; a proxy for stream drying) as having the most explanatory power in predicting benthic macroinvertebrate assemblages in coastal southern California. I contend that if preliminary flow-ecology relationships such as these can be established at streamflow gage sites, they can subsequently be incorporated into an improved hydrologic foundation that explicitly prioritizes the accuracy of the flow metrics that most influence biological endpoints to develop stronger flow-ecology relationships.

The 19 rainfall-runoff models developed for this study were first visually calibrated over the three year period to provide a baseline for subsequent improvement. Next, all 19 models were calibrated to maximize the Nash-Sutcliffe Efficiency. Following this, all 19 models were calibrated to minimize the percent error of the percent of time flow is less than 1 cfs (< 1 cfs). The models were then calibrated to minimize the percent error of the RBI (RBI). Finally, all models were calibrated to simultaneously minimize the percent error of the percent of time when flow is less than 1 cfs and the percent error of RBI based on an arithmetic average of both percent errors (Combined Calibration). Each of the four criteria took approximately the same amount of time and effort to calibrate. Through quantitative calibration, four unique parameter sets were established for each model, producing 76 (19*4) unique models and subsequent hourly time-series of streamflow output for WY 2005-2007. The calibration approaches are summarized in Table 3.5:

Table 3.5: Descriptions of the four calibration criteria.

Calibration Criteria	Flow Regime Component	Description	Best Possible Value	Worst Possibl e Value
NSE	Best overall fit	Uses Eqn. 2.1 to determine the accuracy of each model relative to the observed mean flow during the model duration	One	Negative infinity
< 1 cfs	Low flow	Minimizes the % error of time with flow less than 1 cfs on a time step basis using Eqn. 3.4	0	Infinity
RBI	Streamflow flashiness	Minimizes the % error of RBI using Eqn. 2.2 and Eqn. 3.4	0	Infinity
RBI and < 1 cfs "Combined Calibration"	Flashiness and low flow	Simultaneously minimizes the % error of time with flow less than 1 cfs and RBI using Eqn 2.2 and Eqn 3.4 using a simple average of % errors	0	Infinity

3.4 Streamflow metrics

For each of the 76 modeled time-series of streamflow for WY 2005-2007, IHA 7.1 (Richter *et al.*, 1996) was used to compute 69 descriptive environmental flow component (EFC) metrics. Because IHA operates on a daily time scale, hourly time-series output from HEC-HMS were aggregated to daily. To assess the importance of sub-daily data, NSE, RBI, and the percent of time with flow less than 1 cfs were also calculated at the hourly time scale for all 76 time-series as performance criteria. Often, performance criteria were affected by extreme values and so both the mean and median were computed for each. This same process was repeated for USGS streamflow gage data for the 19 locations. Two sets of gage data, one from WY 2005-2007 and the other from the three year WY period antecedent to bioassessment sampling, which represent recent disturbances affecting individual invertebrate samples, were also used. Bioassessment sampling dates range from 5/8/2001 to 6/12/2012. Descriptions of selected IHA and hourly flow metrics are provided (Table 3.6). Details regarding how these flow metrics were selected from the full suite of 69 IHA flow metrics can be found in Section 3.6.2. For a few sites (11092450 Los Angeles, 11044350 Sandia, 11044300 Santa Margarita Sump), the hourly WY 2005-2007 gage data used for calibration contain no time steps with flow < 1 cfs. This caused calculated < 1 cfs percent errors to occasionally be undefined when a model incorrectly simulates flow < 1 cfs at one of these sites because a value of 0 is used in the denominator. Due to this, the gage and modeled values were switched to compute percent error (Eqn 3.4) when a model incorrectly simulated flow < 1 cfs at a site than contained no flow < 1 cfs.

Table 3.6: Descriptions of hourly and IHA flow metrics.

Flow Metric	Description	Flow Regime Component
RBI_1hr	Richards-Baker Index calculated using Eqn. 2.1 with hourly time steps	Flashiness
cfs_1hr	Percent of time with flow less than 1 cfs using hourly time steps	Low flow
Dec	Median monthly flow	High flow
Jan	Median monthly flow	High flow
Feb	Median monthly flow	High flow
SevenDayMax	Median annual maxima, seven day mean	High flow
NumZeroDays	Median number of zero flow days per year	Low flow
DateMax	Median Julian date of annual one day maximum	High flow
LowPulseCount	Median number of low pulses (50^{th} percentile) within each water year	Low flow
HighPulseCount	Median number of high pulses (75^{th} percentile) within each water year	High flow
RiseRate	Median of all positive differences between consecutive days	Flashiness
FallRate	Median of all negative differences between consecutive days	Flashiness
NumReversals	Median number of hydrologic reversals per year	Flashiness
ExtremeLowDuration	Median annual duration of extreme low flows (10^{th} percentile)	Low flow
ExtremeLowTiming	Median Julian date of extreme low flows (10^{th} percentile)	Low flow
ExtremeLowFreq	Median annual frequency of extreme low flows (10^{th} percentile)	Low flow
HighFlowDuration	Median annual duration of high flows (75^{th} percentile)	High flow
LargeFloodDuration	Median duration of large floods (10 yr return)	High flow
SmallFloodPeak	Median maxima of small floods (2 yr return)	High flow
SmallFloodRiseRate	Median rise rate of small floods (2 yr return)	Flashiness
SmallFloodFallRate	Median fall rate of small floods (2 yr return)	Flashiness
LargeFloodRiseRate	Median rise rate of large floods (10 yr return)	Flashiness

3.5 Benthic macroinvertebrate (biotic) metrics

Taxonomic and trait-based metrics were calculated using SWAMP and SMC sampled benthic macroinvertebrate data. Trait-based metrics are increasingly utilized in stream ecology wherein functional traits, such as resistance to desiccation and other disturbances, are assigned to each taxon and then aggregated into metrics that represent the functional composition of benthic macroinvertebrate assemblages at each site. A large North American database (Poff *et al.*, 2006b) was used to assign traits to taxa. Reach-wide and riffle-specific approaches were used to sample invertebrates at bioassessment sites; however, reach-wide samples were chosen because they were available for more sites than the targeted riffle samples.

For this study, I used a subset of the biological metrics selected by Eberhart (2014) in a preliminary study of streamflow-invertebrate relationships in coastal southern California. The previous study employed correlation, principal component analysis (PCA), literature review, and expert knowledge of flow-ecology relationships (Matt Pyne, Lamar University, personal communication) to reduce nearly 90 taxonomic metrics and over 200 trait-based metrics to four taxonomic metrics, one index of biotic integrity, and seven trait-based metrics. In this present study, only biotic metrics showing strong relationships with flow metrics in the previous flow-ecology study were selected (Table 3.7).

Table 3.7: Descriptions of benthic macroinvertebrate metrics.

Metric Name	Description	Type
EPTPercentTaxa	Invertebrates from orders Ephemeroptera, Plecoptera, and Tricoptera.	Taxa Percent Richness
DesiResist	Organism has one of the following traits advantageous to resistance against desiccation: 1) adult exiting ability 2) desiccation resistance 3) air breather 4) burrowing habit 5) warm eurytherm	Trait Percent Richness
NoninsectTaxa	-	Taxa Richness
DisturbResil	Organism has one of the following traits advantageous to resilience against disturbance: 1) multivoltine 2) fast seasonal development 3) long adult life span 4) strong flying ability 5) high adult female dispersal	Trait Percent Richness
AmphipodaPercent	-	Taxa Percent Abundance
SCIBI	Southern California Index of Biotic Integrity (Ode *et al*., 2005)	Value
SndInstabResist	Organism has one of the following traits advantageous to resistance against bed mobilization in sand-bed systems: 1) burrowing habit 2) sprawling habit 3) streamlined shape 4) adult exiting ability	Trait Percent Richness
ShredderPercentTaxa	Organism in shredder functional feeding group	Trait Percent Richness

3.6 Statistical analyses

All statistical analyses were performed using Microsoft Excel 2010 and 2013, SAS University Edition software (SAS Institute Inc., 2015), and R software version 3.2.0 (package="randomForest"; Liaw and Wiener, 2002). Approaches including percent error, percent difference, correlation, multiple regression (MRA), and random forest (RF) models were used to explore modeled flow metric accuracy and associations between flow and biotic metrics. Percent error was calculated using Equation 3.4:

$$Percent\ Error = \frac{|Gage\ Value - Modeled\ Value|}{|Gage\ Value|} * 100 \qquad \textbf{Eqn. 3.4}$$

Percent difference was calculating using Equation 3.5:

$$Percent\ Difference = \frac{|First\ Value - Second\ Value|}{(First\ Value + Second\ Value)/2} * 100 \qquad \textbf{Eqn. 3.5}$$

Modeled flow metric accuracy and explanatory power were compared to gage data for the period of simulation (WY 2005-2007) and the three WY period antecedent to bioassessment sampling. This approach allows for an investigation into the utility, both in terms of raw flow metric accuracy and explaining variation in benthic macroinvertebrate metrics, of the sub-daily WY 2005-2007 model predictions.

3.6.1 Performance of rainfall-runoff models using flow metrics

The percent error of modeled IHA EFC flow metrics and hourly RBI and percent of time with flow less than 1 cfs metrics were calculated in Microsoft Excel. Modeled metrics were compared to gage data for the WY 2005-2007 calibration period to determine the accuracy of the 19 HEC-HMS models. Percent differences between gage data for the 2005-2007 hydrologic foundation and the three WYs antecedent to each bioassessment sampling were also computed for all sites. Comparison between the two periods of gage data provides insight on how well a constant 2005-2007 simulation period represents antecedent flow conditions experienced by sampled organisms. Correlation analyses were performed in SAS University Edition software using “proc corr” to assess the relationships between flow metrics from all four calibration approaches and both groups of gage data (SAS Institute Inc., 2015).

3.6.2 Performance of rainfall-runoff models using benthic macroinvertebrate metrics

Biotic metrics (Table 3.7) were used to compare how flow metrics generated by the four calibration approaches explain biological variation relative to gage data. Only 14 of the 19 calibrated models were used for these analyses due to the absence of processed bioassessment data at five sites. Using “proc corr” in SAS (SAS Institute Inc., 2015), a correlation analysis explored which of the 71 flow metrics (69 EFCs from IHA, RBI_1hr, and cfs_1hr) computed from gage data for the three WYs preceding each bioassessment sample appear to have strong relationships with each biological metric. The variability of each biotic metric explained by the three WY antecedent gage flow data was used as a comparative standard for statistical analyses because gage flow metrics for the period antecedent to bioassessment sampling might be expected to have a stronger predictive relationship with biotic metrics compared to a constant period.

For each biotic metric, flow metrics from the three WY antecedent gage data significantly correlated at the $p = 0.1$ level were carried forward to a reduced group of predictive flow metrics. Because hourly data have been shown to better represent urbanization and streamflow flashiness which are fundamental to the hydrologic foundation explored in this study, the RBI_1hr and cfs_1hr metrics were included in every reduced group. For some biotic metrics (SCIBI, DisturbResil, SndInstabResist, and ShredderPercentTaxa), less than three flow metrics other than RBI_1hr and cfs_1hr were significantly correlated at the $p = 0.1$ level. For these biotic metrics, the next most significant flow metrics were added to each reduced group so that at least three flow metrics other than RBI_1hr and cfs_1hr were in each biotic metric's reduced group. Each subset of flow metrics from the original 71 metrics described above and unique to every biotic metric will be referred to as the "reduced group of flow metrics" for subsequent analyses. Descriptions of all reduced flow metrics for all biotic metrics are found in Table 3.6.

Using the reduced group of flow metrics unique to each biotic metric, multiple regression analysis was performed with SAS University Edition software separately for each of the four calibration types and two sets of gage data with "proc reg" (SAS Institute Inc., 2015). The entire reduced group of flow metrics for each biotic metric, along with backward selection ($p = 0.05$), minimum C_p selection (Mallows, 1973), and minimum AIC selection (Akaike, 1973) were used in MRA. To further reduce the number of flow metric predictors, the same regression analyses were performed using only the three most significant predictor flow metrics per biotic metric from the three WY antecedent gage data correlation analysis. Standard regression diagnostics were performed and appropriate transformations were made based on visual inspection of residual plots, RStudent residual plots, and Quantile-Quantile plots. To achieve homoscedasticity, AmphipodaPercent was transformed on the log scale along with a couple of the flow metrics from its reduced group: RiseRate and LargeFloodRiseRate. Table 3.8 depicts the reduced group of predictive flow metrics for each biotic metric and indicates which three flow metrics are most significantly correlated with each biotic metric.

Table 3.8: Set of reduced flow metrics for each biotic metric wherein bold flow metrics are the three most significantly correlated.

Biotic Metric	Set of Reduced Flow Metrics
EPTPercentTaxa	RBI_1hr, cfs_1hr, **NumZeroDays**, NumReversals, **ExtremeLowDuration**, **ExtremeLowTiming**, HighFlowDuration, LargeFloodDuration
DesiResist	RBI_1hr, **cfs_1hr**, **NumZeroDays**, LowPulseCount, NumReversals, ExtremeLowDuration, **ExtremeLowTiming**, ExtremeLowFreq, HighFlowDuration, SmallFloodRiseRate, LargeFloodDuration
NoninsectTaxa	RBI_1hr, **cfs_1hr**, **NumZeroDays**, DateMax, **ExtremeLowTiming**, LargeFloodDuration
DisturbResil	RBI_1hr, cfs_1hr, **ExtremeLowTiming**, **LargeFloodDuration**, **HighFlowDuration**
AmphipodaPercent	RBI_1hr, cfs_1hr, **Feb**, **RiseRate**, **ExtremeLowDuration**, FallRate, LargeFloodRiseRate
SCIBI	RBI_1hr, cfs_1hr, **Dec**, **Jan**, **SmallFloodFallrate**
SndInstabResist	RBI_1hr, cfs_1hr, **NumZeroDays**, **ExtremeLowDuration**, **ExtremeLowTiming**
ShredderPercentTaxa	**RBI_1hr**, cfs_1hr, **SmallFloodRiserate**, SevenDayMax, **HighPulseCount**, SmallFloodPeak

A final analysis was performed in R software version 3.2.0 using the "randomForest" package (Liaw and Wiener, 2002). RF models with 10,000 trees were generated using the same list of reduced flow metrics unique to each biotic metric. Similar to the multiple regression analysis, random forest models were used to examine the extent to which variability in biotic metrics could be explained by flow metrics from one of the model calibrations. Likewise, flow metrics produced by gage data were used to explain biotic variability. This was done for every biotic metric, all four calibrations, and both sets of gage data. These analyses allow for comparison of the ability of each calibrated model to explain biotic variability relative to gage data.

CHAPTER 4: RESULTS

HEC-HMS models calibrated to NSE and the three ecologically-focused criteria for the 19 basins showed varying degrees of accuracy (Table 4.1). Unsurprisingly, models reproduced the metrics to which they were calibrated much more accurately than additional flow metrics. RBI calibrated models produced the smallest percent error (median 0.035%), while the Combined Calibration approach produced very low errors (medians less than 1%) despite yielding the lowest NSE (-4.7; Table 4.1). Median percent errors for the three ecologically-focused criteria were all less than 1.5%. Overall, the entire suite of hourly and IHA flow metrics was not very accurately reproduced by the four calibration approaches (Table 4.2), but the < 1 cfs and NSE calibrated models yielded the lowest median percent errors with 54%. Despite a median percent error greater than 100% for the collection of flow metrics produced by the Combined Calibration, some ecologically relevant flow metrics were modeled with accuracy of less than 20% error.

In comparing all gage and modeled flow metrics specific to each biotic metric with MRA of biotic metrics, WY 2005-2007 gage and the Combined Calibration metrics explained slightly more variation in biotic metrics, on average, than three WY antecedent metrics, while the two individual RBI and < 1 cfs ecologically-focused and NSE calibrated models explained more biotic variance in as many MRA models as they explained less (Appendix V). Concerning calibration approaches, the Combined Calibration explained the most biotic variability in 50% of MRA models and never explained the least amount (Table 4.4). Similar results are found for flow metrics selected by minimum AIC through MRA (Table 4.5; Appendix V), and in random forest models (Table 4.6; Appendix VI), but with the NSE and < 1 cfs models explaining the same or slightly more biotic variability, and the RBI models explaining slightly less. In AIC selected MRA results, the Combined Calibration yielded the largest average adjusted R^2 (Table 4.5), indicating it produced the strongest flow-ecology relationships. In RF models, flow metrics from all calibrated HEC-HMS models and the WY 2005-2007 data were able to explain more biotic variability, on average, than the three year antecedent data (Table 4.6).

4.1 Calibration of rainfall-runoff models

The vast majority of the 19 HEC-HMS models were accurately calibrated using the four calibration criteria (Table 4.1). For the three ecologically-focused calibrations (RBI, < 1 cfs, Combined), the specifically targeted ecological flow metric(s) used as calibration criteria were always accurately modeled, with only one model producing its calibration criteria with a greater than 10% error (11.6% for < 1 cfs at 1104250 Rainbow) and median percent errors always less than 1.5%.

Table 4.1: Calibration results displaying NSE, RBI, and < 1 cfs as calibration criteria and performance measures for all 19 models.

		NSE Calibration			RBI Calibration			< 1 cfs Calibration			RBI and < 1 cfs Combined Calibration		
USGS Gage	HEC-HMS Model Name	NSE	RBI % error	< 1 cfs % error	NSE	RBI % error	< 1 cfs % error	NSE	RBI % error	< 1 cfs % error	NSE	RBI % error	< 1 cfs % error
10259000	Andreas	0.575	42.5	100.0	-0.391	0.065	280.1	0.013	99.6	0.3	-4.654	1.9	0.9
11098000	Arroyo Seco	0.579	20.6	271.3	0.575	0.029	271.2	0.150	87.3	0.6	0.457	0.4	0.2
11047300	Arroyo Trabuco	0.753	35.1	648.5	0.172	0.035	1209.9	0.427	72.8	5.1	-0.405	4.5	0.4
11012500	Campo	0.464	72.0	3.0	-2.819	0.175	10.8	0.452	65.8	2.4	-2213.758	0.3	0.7
11044800	De Luz	0.770	43.6	51.6	0.000	0.000	102.2	0.167	97.9	5.5	-0.522	0.3	0.6
10260500	Deep Creek	0.289	18.5	100.0	-0.022	0.096	1802.7	0.287	13.6	1.0	-0.025	4.6	0.1
11015000	Descanso Sweetwater	0.819	60.3	116.8	-0.303	0.004	151.4	0.613	83.5	1.4	-33.264	4.2	0.7
11014000	Jamul	0.593	63.0	100.0	0.266	0.006	11.9	0.498	74.2	1.8	-29.674	0.8	3.0
11092450	Los Angeles	0.885	35.0	0.0	0.553	0.056	100.0	0.885	35.0	0.0	0.609	7.3	0.0
11022200	Los Coches	0.707	8.7	4.8	0.336	0.023	19.2	-2.334	36.6	4.1	-29.280	0.0	0.5
11023340	Poway	0.793	10.4	100.0	0.773	0.044	1103.7	0.794	9.1	3.9	-0.165	5.4	0.3
11044250	Rainbow	0.634	28.6	37.9	0.611	0.029	40.0	0.451	41.2	11.6	0.184	6.6	6.0
11070500	San Jacinto	0.286	91.4	74.0	-60.116	0.074	62.5	-0.098	179.2	1.4	-98.147	0.1	0.8
11042000	San Luis Rey	0.718	8.5	100.0	0.715	0.029	2767.0	-0.412	10.1	7.6	-7.531	0.6	0.7
11046300	San Mateo	0.815	7.1	55.5	0.814	0.076	56.2	0.147	97.7	5.8	-2.997	0.8	4.9
11044350	Sandia	0.854	8.8	100.0	0.839	0.036	100.0	0.606	12.6	0.0	0.312	0.1	0.0
11044300	Santa Margarita Sump	0.578	43.8	0.0	0.129	0.001	0.0	0.578	43.8	0.0	-101.600	6.4	0.0
11044000	Santa Margarita Temecula	0.324	43.5	6815.0	0.319	0.001	5184.9	-10.598	83.3	0.5	-48.079	5.5	0.5
11025500	Santa Ysabel	0.626	17.5	10.4	0.295	0.090	24.2	0.590	13.1	0.5	-56.257	0.4	5.0
	Mean	**0.635**	**34.7**	**457.3**	**-3.013**	**0.046**	**770.5**	**-0.357**	**60.9**	**2.8**	**-138.147**	**2.6**	**1.4**
	Median	**0.634**	**35.0**	**87.0**	**0.295**	**0.035**	**102.2**	**0.427**	**65.8**	**1.4**	**-4.654**	**0.8**	**0.6**

NSE calibrated models produced a median NSE of 0.634. RBI was fairly accurately modeled with the NSE calibrations, producing a median percent error of 35% with seven of the 19 models producing RBI percent errors less than 20%. NSE calibrated models performed poorly when assessed by accuracy of the < 1 cfs metric with a median percent error of 87%, and a maximum of 6,815%. For the NSE calibration criterion, models had final NSE values greater than 0.5 for 15 of the 19 models. Substantial effort went into improving the four models with NSE values less than 0.5 but further improvement was not feasible.

The RBI calibrated models were the most accurate in simulating their calibration criterion with all RBI percent errors less than 0.2% and a median of 0.035%; however, the percent time < 1 cfs was modeled poorly with a median percent error of 102% and a maximum of 5,185%. Nevertheless, a few models (11012500 Campo, 11014000 Jamul, and 11044300 Santa Margarita Sump) produced accurate predictions of percent time < 1 cfs with percent errors less than 20%. Negative NSE values indicate poor model performance for five of the 19 RBI calibrated models, with one outlier producing a NSE of -60. The median NSE of RBI calibrated models was 0.295.

The < 1 cfs calibrated models accurately simulated the percent time < 1 cfs with a median error of 1.4% and only one model (11044250 Rainbow) producing a < 1 cfs percent error greater than 10%. These models yielded a median NSE of 0.427; however, 11044000 Santa Margarita Temecula performed very poorly with an NSE of -10.6 and depresses the mean. Only four of the 19 models calibrated to the < 1 cfs criterion produced negative NSE. RBI was moderately accurately modeled with the < 1 cfs calibrated models, where a median percent error of 66% and a maximum of 179% were produced. The < 1 cfs calibrations, on average, appear to provide the highest overall accuracy when judged by all three metrics.

The RBI and < 1 cfs Combined Calibration proved successful at minimizing the percent error of both metrics. The median RBI percent error equals 0.8% with a maximum of 7.3%, while the median < 1 cfs percent error equals 0.6% with a maximum of 6.0%. Interestingly, NSE was modeled the poorest in the Combined Calibration with a median NSE of -4.7 and a minimum of -2,213.8. Twelve of the 19 models from the Combined Calibration produced negative NSE. For some models and the overall mean and median, the < 1 cfs percent error was actually improved when simultaneously calibrated for RBI and < 1 cfs as opposed to a sole < 1 cfs calibration, likely do to the progression of calibration (Table 3.5).

4.2 Performance of rainfall-runoff models using flow metrics

The mean and median percent errors of IHA and hourly flow metrics relative to the WY 2005-2007 gage data are provided in Table 4.2. The percent differences between flow metrics produced by the three year antecedent gage data and the WY 2005-2007 hydrologic foundation gage data are also included. An error assessment of modeled flows versus the three WY antecedent gage data is provided in Appendix I. In the “Flow Metric” column (Table 4.2; Appendix I), flashiness metrics are denoted in the darker shading and bold, while low flow metrics are lightly shaded and underlined. Additionally, correlation matrices of the entire set of reduced flow metrics produced by data from the two gage sources and four calibrated models are provided in Appendix I.

Table 4.2: Accuracy of flow metrics from three WY antecedent gage data and rainfall-runoff models compared to the calibration period WY 2005-2007 gage data. Italicized average metric errors at the bottom indicate the calibration approach with the lowest average metric error.

Flow Metric	Mean % difference Gage three WY antecedent	Median % difference Gage three WY antecedent	Mean % error NSE Calibration	Median % error NSE Calibration	Mean % error RBI Calibration	Median % error RBI Calibration	Mean % error < 1 cfs Calibration	Median % error < 1 cfs Calibration	Mean % error RBI and < 1 cfs Combined Calibration	Median % error RBI and < 1 cfs Combined Calibration
RBI_1hr	24.1	22.4	15.3	17.6	0.0	0.1	33.6	61.9	0.3	4.6
cfs_1hr	1.0	23.9	23.7	94.0	125.5	262.4	1.1	0.6	0.2	0.2
Dec	34.1	98.5	27.5	41.2	93.7	100.0	223.0	22.9	565.9	336.2
Jan	25.7	84.2	43.8	54.5	80.4	100.0	243.6	59.1	747.8	303.7
Feb	80.5	110.0	39.1	53.7	64.0	100.0	270.3	26.3	819.9	445.6
SevenDayMax	61.2	99.6	38.6	89.5	113.6	82.7	68.0	60.1	646.4	829.2
NumZeroDays	59.8	100.0	135.1	100.0	513.4	100.0	42.5	100.0	117.3	100.0
DateMax	30.0	6.1	54.2	13.7	12.9	5.9	42.2	13.7	16.3	2.0
LowPulseCount	11.5	0.0	82.6	100.0	100.0	100.0	65.1	75.0	27.9	25.0
HighPulseCount	10.1	18.2	7.7	0.0	53.4	100.0	25.9	0.0	86.7	133.3
RiseRate	52.8	40.0	10062.4	1904.5	9451.9	10985.0	5615.0	1658.5	31134.3	38330.0
FallRate	46.7	50.0	434.3	40.8	1483.1	1043.9	209.8	71.3	414.5	518.9
NumReversals	14.3	29.5	72.3	82.9	68.4	76.2	63.3	74.3	39.4	42.9
ExtremeLowDuration	27.3	0.0	19.4	166.7	158.2	6.7	60.7	213.3	64.0	13.3
ExtremeLowTiming	11.2	9.0	29.6	50.0	50.8	74.2	5.3	8.9	14.3	1.6
ExtremeLowFreq	14.6	40.0	10.8	50.0	179.0	200.0	6.2	50.0	47.7	50.0
HighFlowDuration	58.5	40.0	35.4	0.0	32.3	50.5	51.6	25.0	3.0	0.0
LargeFloodDuration	66.1	71.9	17.5	40.8	64.1	83.4	18.1	49.0	69.7	83.4
SmallFloodPeak	23.3	42.6	17.9	54.0	153.3	2.4	73.1	66.2	576.6	671.6
SmallFloodRiseRate	35.0	55.4	24.9	71.5	85.2	39.6	111.5	57.8	318.5	526.4
SmallFloodFallRate	55.8	28.4	62.7	26.7	105.9	290.3	144.4	22.4	204.9	748.9
LargeFloodRiseRate	99.0	106.9	133.8	120.7	137.9	198.8	25.3	11.0	1054.7	559.0
All Low Flow	20.9	16.5	50.2	97.0	187.8	100.0	30.1	62.5	45.2	*19.2*
All Flashiness	46.8	40.0	1543.7	71.5	1618.9	198.8	886.1	*61.9*	4738.1	526.4
All	38.3	40.0	517.7	53.9	596.7	100.0	336.3	*53.9*	1680.5	116.7

In terms of raw flow metric accuracy, the NSE and < 1 cfs calibrated models performed the best overall with median percent errors of 54% for both (Table 4.2). Percent differences between the WY 2005-2007 and three year antecedent gage data were smaller than any model percent error with a median 40% difference. The RBI and Combined calibrations yielded median percent errors greater than 100% across all metrics and generally did not accurately replicate IHA flow metrics. Despite poor overall performance, the Combined Calibration models estimated some IHA flow metrics accurately with less than 20% median error (e.g., ExtremeLowDuration, ExtremeLowTiming, HighFlowDuration, and DateMax; Table 4.2). The Combined Calibration also produced the smallest median percent error of modeled low flow IHA metrics with a value of 19%. Metrics describing rise rates and fall rates at the aggregated daily time step used by IHA were very inaccurate in the Combined Calibration (median percent errors greater than 500%), which affected overall flow metric accuracy. Similar modeled flow metric accuracy was found using the three year antecedent gage data for comparison (Appendix I), but with increased accuracy of the Combined Calibration, especially in regard to NumZeroDays (mean 17% error) and ExtremeLowFreq (median 0% error).

4.3 Performance of rainfall-runoff models based on benthic macroinvertebrate metrics

Results presented in this section will be restricted to summaries of analyses performed for all eight biotic metrics. Results for the individual biotic metrics are provided in Appendices II, III, IV, V, and VI.

4.3.1 *Correlation analysis*

A summary of all biotic metrics correlations with their individual group of reduced flow metrics based on data source is provided (Table 4.3). Appendix II contains correlation tables for each biotic metric individually with its set of reduced flow metrics. Opposing correlation directions indicate when the correlation sign (positive or negative) between biotic and flow metrics were different for modeled or WY 2005-2007 gage data compared to the three year antecedent data.

Table 4.3: Summary of all biotic metrics correlations with unique group of reduced flow metrics per biotic metric.

	3 WY Antecedent Gage	WY 2005-2007 Gage	NSE Calibration	RBI Calibration	< 1 cfs Calibration	RBI and < 1 cfs Combined Calibration
significant correlations	36	22	13	8	10	21
% significant correlations	68%	42%	25%	15%	19%	40%
differing direction of correlation	-	5	25	16	13	11
% differing direction of correlation	-	9.4%	47%	31%	25%	21%
significant differing direction of correlation	-	2	12	1	3	1
% significant differing direction correlation	-	3.8%	23%	1.9%	5.7%	1.9%

As aforementioned, flow metrics were selected for each biotic metric based on significant correlations with the three WY antecedent gage data due to the physical disturbance and habitat relationships between antecedent flow conditions and benthic macroinvertebrates. Because of this, it is unsurprising that the preceding gage data produced the most significant correlations with biotic metrics (68%; Table 4.3). Gage data from the WY 2005-2007 hydrologic foundation period yield the second most significant correlations (42%) and have the least number of differing directions of correlation (9%). The Combined Calibration performed better than any other set of calibrated models with 40% significant correlations and the least number of differing directions of correlation (21%). Flow metrics produced by the NSE calibrated models were oppositely correlated with biotic metrics according to antecedent data nearly half the time (47%).

4.3.2 Multiple regression analysis

MRA was used to examine associations between biotic metrics and flow metrics derived from HEC-HMS models with varying calibration criteria and two sources of gage data. Full MRA results including AIC and adjusted R^2 values are reported for the eight modeled biotic metrics individually and summarized (Table 4.4; Table 4.5; Appendices III-V).

Table 4.4: Multiple regression summary indicating the percent of MRA models in which each set of modeled flow data explained biotic variance the most, second most, and least for both the entire set of reduced flow metrics and the further reduced set of 3 flow metrics. AIC selected flow metrics from both are also included.

Flow Data	**Entire reduced set and set of 3 flow metrics**			**AIC selected flow metrics**		
	Most	Second Most	Least	Most	Second Most	Least
NSE	38%	19%	31%	44%	6%	38%
RBI	13%	25%	38%	13%	25%	44%
< 1 cfs	0%	13%	31%	0%	31%	19%
RBI and < 1 cfs	50%	44%	0%	44%	38%	0%

Table 4.5: MRA results for all eight biotic metrics for both sources of gage data and all four model calibrations using minimum AIC criteria wherein italicized adjusted R^2 values indicate more explanatory power than the three WY antecedent gage data and underlined adjusted R^2 values indicate less explanatory power.

Biotic Metric	3 WY antecedent gage		WY 2005-2007 gage		NSE		RBI		< 1 cfs		RBI and < 1 cfs	
	Adj. R^2	Selected Predictors	Adj. R^2	Selected Predictors	Adj. R^2	Selected Predictors	Adj. R^2	Selected Predictors	Adj. R^2	Selected Predictors	Adj. R^2	Selected Predictors
EPTPercentTaxa	0.32	ExtremeLowDuration, ExtremeLowTiming	*0.79*	RBI_1hr, cfs_1hr, NumZeroDays, NumReversals, ExtremeLowDuration, LargeFloodDuration, HighFlowDuration	*0.52*	RBI_1hr, NumZeroDays, NumReversals, ExtremeLowDuration, ExtremeLowTiming, HighFlowDuration	0.13	NumReversals, HighFlowDuration	0.10	NumZeroDays	*0.49*	cfs_1hr, NumReversals, HighFlowDuration
DesiResist	0.48	ExtremeLowDuration, ExtremeLowTiming	*0.60*	cfs_1hr, ExtremeLowDuration, ExtremeLowFreq, HighFlowDuration	*0.66*	RBI_1hr, LowPulseCount, ExtremeLowFreq, HighFlowDuration, LargeFloodDuration	0.47	RBI_1hr, cfs_1hr, NumZeroDays, NumReversals, ExtremeLowDuration, ExtremeLowTiming, ExtremeLowFreq, SmallFloodRiserate, LargeFloodDuration	*0.59*	RBI_1hr, NumZeroDays, ExtremeLowDuration, HighFlowDuration, SmallFloodRiserate	*0.57*	NumReversals, ExtremeLowDuration, SmallFloodRiserate
NoninsectTaxa	0.36	RBI_1hr, DateMax, LargeFloodDuration	0.31	NumZeroDays, DateMax	0.00	NumZeroDays	*0.47*	RBI_1hr, cfs_1hr, NumZeroDays, DateMax, ExtremeLowTiming	*0.40*	RBI_1hr, NumZeroDays, DateMax, ExtremeLowTiming	0.31	DateMax, ExtremeLowTiming
DisturbResil	0.29	ExtremeLowTiming	*0.36*	ExtremeLowTiming, HighFlowDuration	*0.29*	cfs_1hr, HighFlowDuration	0.04	RBI_1hr, HighFlowDuration	0.02	cfs_1hr	0.21	cfs_1hr, ExtremeLowTiming
Log(AmphipodaPercent)	0.28	log(LargeFloodRiseRate)	*0.38*	cfs_1hr, Feb, FallRate	0.17	RBI_1hr, cfs_1hr	*0.50*	log(RiseRate), ExtremeLowDuration, FallRate, log(LargeFloodRiseRate)	*0.43*	cfs_1hr, log(RiseRate), FallRate, log(LargeFloodRiseRate)	*0.59*	RBI_1hr, Feb, FallRate
SCIBI	0.08	SmallFloodFallrate	*0.10*	Jan	*0.28*	cfs_1hr, Dec, Jan	*0.10*	SmallFloodFallrate	*0.12*	Dec, Jan	*0.35*	RBI_1hr, cfs_1hr, Dec, Jan
SndInstabResist	0.15	ExtremeLowTiming	0.08	ExtremeLowTiming	*0.35*	NumZeroDays, ExtremeLowTiming	0.01	ExtremeLowDuration	0.07	RBI_1hr, NumZeroDays	*0.17*	cfs_1hr, ExtremeLowTiming
ShredderPercentTaxa	0.24	RBI_1hr	0.13	RBI_1hr	0.08	SevenDayMax, SmallFloodPeak	0.13	RBI_1hr	0.09	HighPulseCount	0.18	RBI_1hr, SmallFloodPeak
Mean	0.28		*0.34*		*0.29*		0.23		0.23		*0.36*	
Median	0.29		*0.34*		*0.29*		0.13		0.11		*0.33*	

MRA results using the entire set of flow metrics unique to each biotic metric and only three predictive flow metrics show the Combined Calibration explained the most and second most biotic variability (50% and 44%, respectively) more often than any other calibration approach (Table 4.4). Although the NSE calibrated models explained the most and second most biotic variability more than the other two ecologically-focused calibrations, it often explained the least variability (31%). The Combined Calibration approach never explained the least biotic variability. The < 1 cfs calibration provided the highest overall accuracy for raw flow metrics (Table 4.2); however, it never explained the most biotic variability.

Flow metrics selected using minimum AIC from the WY 2005-2007 gage, NSE calibrated, and Combined Calibration data explain as much or more biotic variability, on average, compared to the antecedent gage data (Table 4.5). The Combined Calibration explained the most biotic variability across metrics (mean adj. R^2 = 0.36), while the NSE calibrated data explained slightly more biotic variability, on average, than the antecedent gage data. The NSE data yielded low adjusted R^2 values for a few metrics (0 for NoninsectTaxa, 0.17 for log(AmphipodaPercent), and 0.08 for ShredderPercentTaxa), while the Combined Calibration produced its smallest adjusted R^2 of 0.18 for ShredderPercentTaxa. The two individual RBI and < 1 cfs ecologically-focused calibration approaches explained the least amount of biotic variability, on average, in AIC selected MRA.

In full MRA analyses, WY 2005-2007 gage and the modeled Combined Calibration metrics explained slightly more variation in biotic metrics, on average, than three WY antecedent metrics when considering the entire group of flow metrics unique to each biotic metric (Appendix III). Flow metrics from the other two ecologically-focused and NSE calibrated models explained biotic variance equally as well as flow metrics from the three WY antecedent data, on average (Appendix III). When the number of predictive flow metrics was reduced to three, biotic variation explanation also decreased; however, the WY 2005-2007 gage and Combined Calibration data still explained more variance than the NSE, RBI, or < 1 cfs calibrated models.

4.3.3 Random forest analysis

RF was also used to compare calibration approaches and gage data by examining relationships between biotic and flow metrics. Full RF results for the eight modeled biotic metrics individually and summarized are provided (Table 4.6; Table 4.7; Appendix VI).

Table 4.6: Random forest summary indicating the percent of RF models in which each set of modeled flow data explained biotic variance the most, second most, and worst for the set of reduced flow metrics.

Flow Data	**Entire reduced set of flow metrics**		
	Most	Second Most	Least
NSE	13%	50%	25%
RBI	13%	13%	63%
< 1 cfs	38%	13%	0%
RBI and < 1 cfs	38%	25%	13%

Table 4.7: Random forest results for all biotic metrics with each set of reduced flow metrics. Percent variance values for flow data are italicized when more biotic variance is explained than the 3 WY antecedent gage data and percent variance values for flow data are underlined when less is explained.

Biotic Metric	3 WY antecedent gage		WY 2005-2007 gage		NSE		RBI		< 1 cfs		RBI and < 1 cfs	
	% Var.	Important Predictors	% Var.	Important Predictors	% Var.	Important Predictors	% Var.	Important Predictors	% Var.	Important Predictors	% Var.	Important Predictors
EPTPercent Taxa	9.6	ExtremeLowTiming, LargeFloodDuration, NumReversals	*34.0*	HighFlowDuration, NumReversals, ExtremeLowTiming	*33.4*	ExtremeLowTiming, ExtremeLowDuratio n, NumReversals	*13.9*	NumReversals, ExtremeLowTiming, ExtremeLowDuration	*27.9*	RBI_1hr, ExtremeLowTiming, HighFlowDuration	*34.2*	NumReversals, ExtremeLowDuration, LargeFloodDuration
DesiResist	11.2	ExtremeLowTiming, NumReversals, LargeFloodDuration	*29.4*	SmallFloodRiseRate, NumReversals, HighFlowDuration	*28.7*	ExtremeLowTiming, NumReversals, ExtremeLowFreq	*16.0*	NumReversals, LargeFloodDuration, cfs_1hr	*24.5*	cfs_1hr, ExtremeLowTiming, ExtremeLowFreq	*29.6*	NumReversals, ExtremeLowTiming, ExtremeLowDuration
NoninsectTa xa	0 (-)	ExtremeLowTiming, RBI_1hr, cfs_1hr	*15.2*	LargeFloodDuration, DateMax, cfs_1hr	*0 (-)*	ExtremeLowTiming, LargeFloodDuration, RBI_1hr	*0 (-)*	NumZeroDays, cfs_1hr, RBI_1hr	*10.8*	cfs_1hr, DateMax, ExtremeLowTiming	*4.5*	cfs_1hr, LargeFloodDuraton, ExtremeLowTiming
DisturbResil	0 (-)	ExtremeLowTiming, LargeFloodDuration, cfs_1hr	*19.9*	HighFlowDuration, ExtremeLowTiming, cfs_1hr	*0.2*	cfs_1hr, ExtremeLowTiming, LargeFloodDuration	<u>0 (-)</u>	HighFlowDuration, cfs_1hr, ExtremeLowTiming	*8.6*	RBI_1hr, LargeFloodDuration, ExtremeLowTiming	<u>0 (-)</u>	LargeFloodDuration, cfs_1hr, ExtremeLowTiming
Log(Amphip odaPercent)	15.9	log(LargeFloodRiseRat e), Feb, ExtremeLowDuration	*20.6*	cfs_1hr, Feb, ExtremeLowDuratio n	*20.2*	Feb, ExtremeLowDuratio n, RBI_1hr	*29.7*	cfs_1hr, log(RiseRate), log(LargeFloodRiseRat e)	*27.2*	Feb, cfs_1hr, log(RiseRate)	*32.8*	cfs_1hr, Feb, FallRate
SCIBI	0 (-)	SmallFloodFallRate, RBI_1hr, Dec	*12.3*	SmallFloodFallRate, RBI_1hr, Dec	*3.3*	SmallFloodFallRate, Dec, Jan	*3.0*	SmallFloodFallRate, RBI_1hr, cfs_1hr	*9.6*	SmallFloodFallRate, Jan, cfs_1hr	*6.6*	RBI_1hr, SmallFloodFallRate, cfs_1hr
SndInstabRe sist	0 (-)	ExtremeLowTiming, cfs_1hr, RBI_1hr	*0 (-)*	ExtremeLowTiming, RBI_1hr, cfs_1hr	*22.7*	ExtremeLowTiming, ExtremeLowDuratio n, RBI_1hr	<u>0 (-)</u>	ExtremeLowDuration, ExtremeLowTiming, RBI_1hr	*4.3*	ExtremeLowTiming, RBI_1hr, cfs_1hr	*1.8*	ExtremeLowTiming, RBI_1hr, ExtremeLowDuration
ShredderPer centTaxa	0 (-)	RBI_1hr, SmallFloodPeak, SmallFloodRiseRate	<u>0 (-)</u>	SmallFloodRiseRate, RBI_1hr, HighPulseCount	<u>0 (-)</u>	RBI_1hr, SevenDayMax, HighPulseCount	<u>0 (-)</u>	RBI_1hr, SevenDayMax, HighPulseCount	<u>0 (-)</u>	HighPulseCount, SmallFloodRiseRate, cfs_1hr	<u>0 (-)</u>	RBI_1hr, SevenDayMax, HighPulseCount
Mean	4.6		*16.4*		*13.6*		*7.83*		*14.1*		*13.7*	

Random forest results were similar, albeit slightly different than MRA. Using the reduced set of flow metrics unique to each biotic metric in a random forest analysis, the < 1 cfs and Combined Calibration approaches explained the most variance in the same number of RF models (37.5%; Table 4.6). Unlike the multiple regression analysis, the Combined Calibration did explain the least variability in 12.5% of random forest models. Comparatively, NSE did not perform well using random forest as it explained the most biotic variability in only 12.5% of models and the least in 25%. RBI performed the worst in random forest models considering the reduced sets of flow metrics.

Overall, RF models explain more biotic variability when data from the four calibrated models and WY 2005-2007 period were used, on average, than when the antecedent gage data was used (Table 4.7). The WY 2005-2007 gage data explained the most biotic variability (mean of 16%), while the < 1 cfs calibration explained more than any other calibration approach (mean of 14%). The NSE and Combined Calibration approaches performed similarly to < 1 cfs, with means slightly less than 14%, and the RBI calibration explained the least variability (mean of 7%).

Appendix VI provides a table indicating how often more or less biotic variability was explained by the four types of calibrated models and WY 2005-2007 gage data compared to the three WY antecedent gage.

CHAPTER 5: DISCUSSION

5.1 Ecologically-focused calibration of hydrological models

Rainfall-runoff models should be created to solve specific problems that are understood before creating and applying the models (Beven and Binley, 1992), and prior selection of and targeted calibration to specific streamflow characteristics have been recommended to more accurately model ecological flow metrics (Cassin *et al.*, 2005; Murphy *et al*., 2013). This study compares novel, ecologically-focused calibration criteria based on prior knowledge of regional flow-ecology relationships to a traditional best overall fit criterion by modeling flow metrics and assessing the strength of flow-ecology relationships based on those metrics.

Results indicate that predictive flow-ecology relationships produced by rainfall-runoff models calibrated using ecologically-focused criteria based on preliminary regional flow-ecology relationships have potential to provide a more useful hydrologic foundation for ecohydrological studies compared to traditional overall fit calibration criteria. Specifically, metrics produced by the Combined Calibration approach provide the strongest correlations between flow and biotic metrics, explain the most biotic variability in MRA, on average, and perform very well in RF models. In all statistical analyses, the Combined Calibration explains more biotic variation, on average, than the traditional, best overall fit NSE calibration criterion; however, NSE calibrated models explain more biotic variation than the individual RBI and < 1 cfs calibrations in MRA and more biotic variation than the individual RBI calibration in RF, on average.

The Combined Calibration models reproduce the full suite of IHA flow metrics with a low level of accuracy when viewed collectively (median percent error greater than 100%; Table 4.2); however, the Combined Calibration approach predicts some of the most ecologically relevant streamflow metrics with a much higher degree of accuracy. These flow metrics include ExtremeLowTiming, NumZeroDays, cfs_1hr, RBI_1hr, HighFlowDuration, LargeFloodDuration, and ExtremeLowDuration. For the establishment of a regional hydrologic foundation, it is important that hydrologic models accurately simulate hydrologic processes; however, the ultimate aim is to accurately predict the most influential elements of the flow regime that shape biological communities for flow-ecology relationships (Poff *et al.*, 2010; Kendy *et al.*, 2012; TNC, 2015). Hydrological models can never perfectly replicate hydrologic processes and tradeoffs must be made to prioritize the accuracy of the most important elements of the flow regime for specific model applications. Low flow and flashiness metrics typical of ephemeral, Mediterranean streams have historically been difficult to model accurately. As expected, the Combined Calibration is much less accurate for simulating elements of streamflow that appear to be less important predictors of benthic macroinvertebrate metrics; however, the tradeoff appears worthwhile due to substantial biological explanation by the Combined Calibration flow metrics in this study. Furthermore, out of the three ecologically-focused calibration criteria, < 1 cfs produced the highest median NSE and estimated IHA flow metrics with the lowest error, but explained less overall variability in biotic metrics than the Combined Calibration, which produced the lowest median NSE.

Models specifically calibrated to NSE accurately replicated some flow metrics, especially those describing large flow events (HighFlowDuration, LargeFloodDuration, SmallFloodPeak, SmallFloowRiseRate, and SmallFloodFallRate). These high flow metrics were selected as significant predictors of biotic metrics less frequently than low flow or flashiness metrics, but were used in some flow-ecology relationships. The Combined Calibration accurately reproduced only two high flow metrics (DateMax and HighFlowDuration); however, these metrics appeared the most frequently in MRA results which helped strengthen the Combined Calibration results. I do not think the explanatory power of high flow metrics is an anomalous statistical artifact. Instead, some high flow metrics, especially HighFlowDuration, play a role in shaping biological communities. Periods of high flow are biologically important due to their disturbance effects, flushing of fine sediment and general reshaping of aquatic habitats, and facilitation of life cycle completion for aquatic insects and fish. This study focused on optimizing nontraditional but known ecologically relevant aspects of the flow regime unique to arid and Mediterranean climates, such as coastal southern California, that are not emphasized in traditional rainfall-runoff model calibration. This is not to say, however, that high flows are not biologically relevant in coastal southern California. Clearly high flow metrics should not be ignored and are still shown to be biologically relevant, but in unique climatic settings such as coastal southern California prioritizing the accuracy of flow metrics known *a priori* to be especially strong or regionally unique to ecological endpoints may produce stronger flow-ecology relationships.

Data quality is an important consideration that may have influenced some results in this study. Four models were calibrated with NSE values less than 0.500 (10260500 Deep Creek, 11012500 Campo, 11070500 San Jacinto, and 11044000 Santa Margarita Temecula). According to USGS reports for these gages, flows are regulated due to reservoirs, effluent discharge, irrigation diversions, and municipal diversions. These flow alterations often created instances when precipitation occurred but no pulse was seen in gaged streamflow records, or instances when streamflow pulses occurred with no precipitation. These scenarios are impossible to model without a complete water balance tracking all water leaving and entering the stream; however, flow regulation is common but is not systematically documented. Such issues highlight the importance of tradeoffs in hydrological modeling due to a practical inability to represent all hydrologic processes. Furthermore, a relatively small biotic dataset was used in this study (25 sites; 8 metrics), which could have directly influenced the accuracy of flow-ecology results of this study. The use of presence/absence biotic data in this study helps reduce the impact of these errors compared to more sensitive biotic metrics, such as those characterizing abundance. Nevertheless, the flow ecology-relationships developed for ungaged sites should be considered preliminary until the number of sites is appreciably increased.

RF model results deviate slightly from those produced by MRA. A small sample size (n = 25) and intercorrelation between some predictor flow metric variables are not handled well in RF analysis and are likely the cause of discrepancies. Furthermore, RF modeling was used to explain biotic variability in the same manner as MRA, and the traditional approach of identifying thresholds and cutoffs among predictor variables was not implemented. This study facilitated my first attempt experimenting with RF statistical modeling, and so it is included despite its noted weaknesses in this specific application. MRA results should be prioritized over RF results for this study.

The hydrological models developed in this study were created to help guide management decisions for establishing environmental streamflow criteria. Based on the modeled accuracy of ecologically relevant flow metrics and strength of flow-ecology relationships, I think the Combined Calibration models could be reasonably used to make management decisions about stream drying and flashiness aspects of the flow regime. Since the models were calibrated to these two elements of the flow regime specifically, and were intended for environmental streamflow criteria development, I would feel comfortable using them for this specific purpose. However, the models are probably unsuitable for other applications without further calibration and testing.

These findings suggest the importance of utilizing prior knowledge of flow-ecology relationships during calibration and confronting modeled flow metrics with biological data to rigorously assess modeled hydrologic foundations. Previous studies have compared regional regression approaches for generating ecological flow metrics to hydrologic modeling approaches (Murphy *et al.*, 2013) and compared best overall fit calibration criteria for modeling ecological flow metrics (Vis *et al.*, 2015), but I could find no studies that examine the predictive utility of specific flow regime elements as a benchmark for comparison. The findings of this study also suggest that calibration for general fit to a large and arbitrary suite of flow metrics (e.g., the complete set of IHA metrics) may not provide the best information for flow-ecology applications. Biological knowledge should inform the development of a hydrologic foundation because the most ecologically relevant flow metrics may be very accurate despite other inaccuracies. The specific ecologically-focused calibration criteria used in this study were chosen *a priori* based on relationships between regional streamflow gages and biota (Eberhart *et al.,* 2014), suggesting similar preliminary analyses can be used to guide calibration of rainfall-runoff models for hydroecological studies in different regions.

5.2 Inadequacy of Nash-Sutcliffe Efficiency for some rainfall-runoff model applications

Previous research indicates that best overall fit criteria, such as NSE, are not always the ideal calibration criteria to use for rainfall-runoff models (Cassin *et al.,* 2005; Jain and Sudheer, 2008; Beven, 2012); however, the adequacy of using such calibration criteria in developing a hydrologic foundation for flow-ecology relationships within an environmental flow criteria framework has only recently been explored (Vis *et al.*, 2015). Through unique comparisons of a best overall fit calibration criterion with ecologically-focused alternatives, this study bolsters the idea that NSE may be inappropriate for specific model applications, such as the development of environmental flow criteria.

The results from this study indicate that an ecologically-focused Combined Calibration approach can consistently produce models that simulate flow-ecology relationships more accurately than NSE calibrated models, at least for coastal southern California streams. This is supported by multiple lines of evidence. Correlation results show that NSE models are least predictive in relation to the set of biotic metrics I examined (Table 4.3). MRA results indicate that the Combined Calibration outperforms NSE, on average and that, there are a relatively large number of instances in which the NSE calibrated models explained the least biotic variability (Table 4.4). In the random forest analysis, the < 1 cfs calibrated and Combined Calibration models explain more biotic variability than the NSE calibrated models (Table 4.6). These findings and the fact that NSE does not provide easier or quicker calibrations compared to ecologically-focused criteria indicate that NSE should not be blindly applied to calibrating models for hydroecological studies.

Overall, NSE models produced the entire suite of flow metrics with a higher level of average accuracy than the Combined Calibration (median 54% error); however, these flow metrics explained less variability in the selected biotic metrics because those with the most biological explanatory power were less accurately modeled. Despite the Combined Calibration models producing mostly negative NSE values, they explained more biotic variability than models optimizing NSE. Similarly, the < 1 cfs calibrated models produced a larger median NSE value than the ecologically-focused Combined Calibration, but did not explain invertebrate metrics as well. This finding again suggests the importance of confronting modeled flow metrics with biological data early in the hydrologic modeling process if possible to identify ecologically relevant metrics for developing a deeper understanding of hydrologic foundation robustness, and not relying simply on collective raw flow metric accuracy. If biological data were not incorporated in this study, as has been the norm, different interpretations of hydrologic foundation accuracy and flow-ecology relationships would be achieved.

5.3 Calibration of hydrological models using multiple criteria

Based on the superior performance of the Combined Calibration over other calibration attempts, this study bolsters the idea that multiple criteria should be used simultaneously to calibrate rainfall-runoff models. While it is recommended that multiple performance measures be used to assess rainfall-runoff model accuracy (Beven, 2012), criteria that simultaneously incorporate multiple performance measures for different elements of the flow regime are far less commonly used but have shown to produce more robust models (Gupta *et al.*, 2008), albeit outside the context of environmental flow applications.

When developing rainfall-runoff models with the purpose of accurately predicting several flow regime characteristics important to biological endpoints, as necessitated by environmental flow studies, Murphy *et al.* (2013) and Vis *et al.* (2015) recommend multiple calibration criteria focused on overall fit. In this study, the Combined Calibration multiple criteria approach attempted to optimize low flow and flashiness aspects of the flow regime without considering best overall fit. In doing so, other elements of the flow regime, such as magnitude, were not considered and were modeled less accurately. This substantially increased the average metric percent error; however, crucial ecological flow metrics were produced with low percent errors resulting in the Combined Calibration generally providing the most biotic explanation. Because the ultimate aim of rainfall-runoff modeling within a hydrologic foundation is accurate flow-ecology relationships, and not necessarily overall accuracy of all flow metrics, the multiple calibration criteria approach appears superior to a NSE overall fit criterion. This finding should not overshadow the importance of overall fit for accurately identifying flow-ecology relationships which utilize elements of the flow regime not identified *a priori*. While this study demonstrates that ecologically-targeted calibration of hydrological models can improve flow-ecology relationships versus a best overall fit criterion, more sophisticated calibration approaches that simultaneously optimize ecologically-focused and best overall fit criteria may further strengthen flow-ecology relationships not identified prior to modeling.

5.4 Using antecedent gage data as the standard for comparing modeled flow-ecology relationships

Three WY antecedent gage data were used for reducing flow metric sets and as the standard upon which the explanation of biotic variability was compared in this study. Because three years of flow data have been shown to sufficiently explain benthic macroinvertebrate variability (Kennen *et al.*, 2010; Eberhart, 2014), and due to the formidable challenges associated with using hourly data, three year periods of record were considered in this study. Habitat disturbances are a primary physical mechanism through which streamflow directly influences benthic macroinvertebrate assemblages. These disturbances are often caused by modified stream channel form due to bankfull and overbank flooding events which often occur every 1.5 – 2 years (Soar and Thorne, 2001), well within the three years used in this study. The WY 2005-2007 period was specifically chosen due to data availability, and because it represents a typical wet, moderate, and dry year similar to the long-term climate in coastal southern California; however, it does not incorporate exceptionally anomalous years that may occur in the three year antecedent gage data for each bioassessment sample. Because WY 2005-2007 captures the representative range of regional climate variability, it could be used to reduce ecological flow metrics and as the standard for comparing flow-ecology relationships at the expense of encompassing disturbance events that occurred prior to biological sampling outside the 2005-2007 period.

Because results from this study indicate that flow metrics representing the standard period (WY 2005-2007) are generally similar to the three WY antecedent period (mean percent difference of 38% and no individual metric different by more than 100%), it is likely that similar results would be obtained if WY 2005-2007 gage data were used as the standard for flow metric reduction and statistical comparison with biotic metrics. Confrontation of the biological data with two periods of streamflow gage data indicates similarly strong flow-ecology relationships according to correlation, multiple regression, and random forest analyses; however, for every multiple regression and random forest analysis, the WY 2005-2007 gage data explained biotic variability, on average, as well or better than the three WY antecedent gage data. These results are likely due to the similarity of flow metrics produced by the two sources of gage data and the representative range of regional variability captured by the hydrologic foundation period. While benthic macroinvertebrate assemblages are certainly shaped by three year antecedent streamflow, it appears as if "typical" flow metrics produced by the WY 2005-2007 period in this study may more accurately represent differences in relative hydrologic behavior among sites over a common period.

Gage data from three antecedent WYs were poor predictors of substantial biotic variability (0% in RF and adj. $R^2 < 0.15$ in MRA) for some biotic metrics used in this study, especially in the random forest analyses. In many instances, the WY 2005-2007 gage data improved explanatory power, often to positive percent variability explained (RF) and more substantial adjusted R^2 values (MRA). Perhaps for some biotic metrics, anomalous events in the three years preceding sampling have weakened flow-ecology relationships that are better described with a common representative climatic period. Specifically, the biotic metrics used in this study are all based on presence or absence of a specific taxa and not abundance, with the exception of AmphipodaPercent. These presence/absence biotic metrics are likely relatively unaffected by anomalous events in the three WY antecedent to sampling, while abundance biotic metrics are likely more sensitive to anomalous disturbance events. If abundance biotic metrics had been used in this study, the three WY antecedent conditions might have explained more biotic variability. A more systematic comparison of hydrologic foundation durations (3 yr vs longer) with the different types of data used in this study (three year antecedent vs. constant climate) and types of biotic metrics (presence/absence vs. abundance) for coastal southern California could provide additional insight.

Flow metrics that explained the most biotic variability in multiple regression and random forest analyses were commonly not the same between gage and modeled data; however, they each came from a group of similar metrics that was identified *a priori* as ecologically relevant using antecedent gage data. Extensive review of regional literature and, if available, gage data associated with nearby bioassessment sites should be used to guide hydrological model calibration to improve the likelihood that modeled flow-ecology relationships are sufficiently strong to provide guidance on management options.

5.5 A three year duration and hourly resolution hydrologic foundation

Both hourly flow metrics derived from three years of data (the percent of time with flow < 1 cfs and RBI) explained significant variance in the benthic macroinvertebrate metrics examined in this study; however, it is likely that a three year hydrologic foundation with an hourly resolution is appropriate only for benthic macroinvertebrates or other relatively short-lived biota that are sensitive to stream drying events and flashiness. A three year period may not be adequate to identify relationships with more mobile or longer-lived biota, such as fishes or riparian forest vegetation, and it is not likely that hourly flow data will provide any additional benefit over daily for these biota.

The temporal resolution of streamflow data significantly affects modeled flashiness. In this study, hourly time-series were aggregated to produce daily IHA flow metrics. This resolution coarsening appears to have greatly affected the accuracy of the rise and fall rate metrics in calibrated models, which were consistently among the least accurately modeled (Table 4.2) and not commonly selected as significant predictors from both gage and modeled metrics. This is especially true for the Combined Calibration where some rise and fall rates were off by two orders of magnitude. While the Combined Calibration and RBI calibration successfully minimized RBI percent error at the hourly resolution (medians less than 1%), they did not adequately model rise and fall rates at the daily time scale after information content was reduced through aggradation. Sub-daily models accurately calibrated to flashiness may be erroneous when aggregated to daily time steps, as demonstrated by the large errors in daily IHA flashiness flow metrics produced by ecologically-focused models specifically calibrated to optimize the accuracy of streamflow flashiness at hourly time steps; however, sub-daily flashiness metrics may significantly enhance flow-ecology relationships, as demonstrated in this study, due to their biological relevance in certain regions, such as coastal southern California.

In coastal southern California, WYs 2005, 2006, and 2007 are characterized as wetter than average, average, and drier than average, respectively. These years represent the range of the regional climatic variability and it is not likely that their order of occurrence negatively impacted model calibration, especially for the Combined Calibration. The Combined Calibration includes a criterion more sensitive to higher flows relevant to the beginning of the modeled time period (RBI) and a criterion focused on lower flows more relevant to the end of the modeled period (< 1 cfs). In coastal southern California, streamflow flashiness is shaped by large events that produce high pulses of flow for very short periods of time. Furthermore, no systematic residual trends were observed across all models for the calibration approaches more sensitive to high flows (NSE, RBI, Combined) and those sensitive to low flows (< 1 cfs, Combined).

5.6 Future research

Ecologically-focused calibration criteria for hydrological models have shown promise in this study and warrant further investigation. If it proves impossible to accurately model all elements of the flow regime, as is the case with long-term, continuous rainfall-runoff models, calibration criteria relevant to specific biological endpoints, and combinations of these criteria, should be investigated in additional regions for alternative biologic endpoints to determine if more robust hydrologic foundations can be created for ecohydrological studies. Specifically, multi-criteria approaches that simultaneously incorporate ecologically-focused and best overall fit criteria should be explored. A quantitative analysis of the tradeoffs (flow-ecology relationship strength and flow metric accuracy) involved between ecologically-focused and best overall fit calibration schemes could provide insight on when highly sophisticated multiple calibration criteria warrant use. A deeper investigation into the duration, resolution, and type of streamflow data for establishing a hydrologic foundation with benthic macroinvertebrates specifically, across many regions, should also be explored. The adequacy of three years of hourly streamflow data in describing other organisms could be tested by generating flow metrics from time-series data with associated bioassessment sites across gradients of time resolution and duration. Time steps could range from 15 min to daily or monthly and durations could range from three years to at least 15 years. Correlation, MRA, and other statistical procedures could be used to determine the explanatory power of different resolutions and durations of data for various biotic endpoints.

Hydrological models used for environmental flow applications could be used to simulate possible stormwater management scenarios. The impacts of different management approaches generated by ecologically-focused models on local biota could help decision-makers consider ecological health in management choices. Furthermore, alternative ecologically-focused models calibrated to hydraulic conditions may more accurately simulate habitat disturbance events that most affect stream biota. The strength of flow-ecology relationships produced by hydraulic ecologically-focused calibration criteria warrants future research.

Accurate predictions of streamflow at ungaged bioassessment sites are often needed to establish regional flow-ecology relationships. While a substantial number of studies have recently focused on predicting streamflow in ungaged basins (Bloschl *et al.*, 2013), more robust techniques are needed to extrapolate one or a combination of calibrated rainfall-runoff models to ungaged basins. When model parameters cannot be directly estimated, recent research has focused on spatial proximity and similarity in key watershed characteristics as the basis for extrapolation; however, approaches that implement model accuracy through cross-validation should be explored further. Commonly, two neighboring watersheds may generate dramatically different hydrological processes and so the accuracy of cross-validated rainfall-runoff models when extrapolated to other basins within a region could be an important tool for predictions in ungaged basins.

CHAPTER 6: CONCLUSIONS

Nineteen rainfall-runoff models spanning three years with hourly time steps were individually calibrated to streamflow gages in coastal southern California. Calibration criteria were NSE and three ecologically-focused criteria identified *a priori* based on knowledge of regional flow-ecology relationships. Calibration was very accurate with only one model calibrated to ecologically-focused criteria producing a calibration error greater than 10%. Many models were calibrated within a 1% error of their ecologically-focused criteria. Calibration to NSE was successful with a median value of 0.634.

Correlation, multiple regression, and random forest analyses indicate that flow metrics produced by the ecologically-focused Combined Calibration approach generally explain selected taxa and trait-based biotic metrics as well or better than metrics derived from gage data in coastal southern California. Furthermore, these models calibrated simultaneously to two ecologically-focused criteria generally produce superior flow-ecology relationships, on average, compared to those calibrated to the NSE best overall fit criterion. Despite their relatively low accuracy in reproducing some of the full suite of IHA flow metrics, models calibrated based on biologically-relevant criteria appear to reproduce biologically important flow metrics related to extreme low flows and flashiness much more accurately than models calibrated using an overall fit criterion. This can lead to stronger flow-ecology relationships, the ultimate aim of hydrological modeling in an environmental flow framework. Ecohydrological studies that utilize rainfall-runoff models and report calibration methods almost exclusively employ a best overall fit criterion for calibration. Such an approach does not necessarily focus on elements of the flow regime critical to biologic endpoints, and may not always be the preferred option as demonstrated in this study.

This study shows how prior knowledge of flow-ecology linkages can be incorporated into a hydrologic foundation.

A hydrologic foundation spanning three years that typify a wet, moderate, and dry year in coastal southern California, and providing hourly time series of discharge, produces flow-ecology relationships that are, on average, at least as strong as those developed from gage data for the three years preceding each bioassessment sample. Raw flow metrics produced by the gaged constant climate period were more similar to those produced by the three year antecedent data than any modeled data. Based on these findings, the WY 2005-2007 hydrologic foundation period appears to better describe benthic macroinvertebrates in coastal southern California than three year antecedent gage data scattered over eleven years.

None of the hourly models performed exceptionally well in reproducing the full suite of daily IHA metrics with the smallest overall percent error being a median value of 54% (NSE and < 1 cfs). This is not necessarily surprising due to the prevalence of magnitude-based flow metrics in IHA and the effects of aggregating to daily time steps. While the Combined Calibration produced the largest overall median percent error of collective flow metrics (117%), it more accurately reproduced certain flow metrics critical to benthic macroinvertebrate endpoints (medians < 20%).

Future research is needed to investigate how to achieve best overall fit, while also ensuring accuracy of the most ecologically relevant elements of the flow regime. Furthermore, the use of *a priori* flow characteristics known to shape biological endpoints as rainfall-runoff calibration criteria should be expanded to different regions and biota to better assess the feasibility of improving hydrologic foundations within environmental flow studies. This research informs regional studies in which calibrated rainfall-runoff models must be extrapolated to ungaged bioassessment sites for a more inclusive regional hydrologic foundation.

REFERENCES

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. *2nd International Symposium on Information Theory*, Tsahkadsor, Armenia, USSR, September 2-8, 1971, Budapest: Akadémiai Kiadó, p. 267-281.

Baker, D.B., P. Richards, T.T. Loftus, and J.W. Kramer (2004). A new flashiness index: characteristics and applications to Midwestern rivers and streams. *Journal of the American Water Resources Association (JAWRA)* **40**(2):503-522, DOI: 10.1111/j.1752-1688.2004.tb01046.x.

Barbour, M.T., J. Gerritsen, B.D. Snyder, and J.B. Striblin (1999). Rapid Bioassessment Protocols for Use in Strams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish, Second Edition. Report EPA 841-B-99-002., U.S. Environmental Protection Agency, Office of Water, Washington, D.C., URL: http://water.epa.gov/scitech/monitoring/rsl/bioassessment/.

Belmar, O., J. Velasco, C. Gutierrez-Canovas, A. Mellado-Diaz, A. Millan, and P.J. Wood (2013). The influence of natural flow regimes on macroinvertebrate assemblages in a semiarid Mediterranean basin. *Ecohydrology* **6**(3):363-379, DOI: 10.1002/eco.1274.

Bevelhimer, M.S., R.A. McManamay, and B. O'Connor (2014). Characterizing Sub-Daily Flow Regimes: Implications of Hydrologic Resolution on Ecohydrology Studies. *River Research and Applications*, 13 p., DOI: 10.1002/rra.2781.

Beven, K.J. (1989). Changing ideas in hydrology—the case of physically-based models. *Journal of Hydrology* **105**(1-2):157-172, DOI: 10.1016/0022-1694(89)90101-7.

Beven, K.J. (2012). Rainfall-Runoff Modelling: The Primer – Second Edition. John Wiley & Sons, West Sussex, United Kingdom, DOI: 10.1002/9781119951001.

Beven, K.J. and A. Binley (1992). The Future of Distributed Models: Model Calibration and Uncertainty Prediction. *Hydrologic Processes* **6**(3):279-298, DOI: 10.1002/hyp.3360060305.

Bledsoe, B.P. and C.C. Watson (2001). Effects of Urbanization on Channel Instability. *Journal of the American Water Resources Association (JAWRA)*, **37**(2): 255-270, DOI: 10.1111/j.1752-1688.2001.tb00966.x.

Bledsoe, B.P., M.C. Brown, and D.A. Raff (2007). GeoTools: a toolkit for fluvial system analysis. *Journal of the American Water Resources Association (JAWRA)*, **43**(3):757–772, DOI: 10.1111/j.1752-1688.2007.00060.x.

Bloschl, G., M. Sivapalan, T. Wagener, A. Viglone, and H. Savenije (2013). Runoff Prediction in Ungauged Basins: Synthesis across Processes, Places, and Scales. Cambridge University Press, London, United Kingdom, DOI: 10.1002/2014EO020025.

Booth, D.B. (1990). Stream-Channel Incision Following Drainage-Basin Urbanization. *Journal of the American Water Resources Association (JAWRA)* **26**(3):407-417, DOI: 10.1111/j.1752-1688.1990.tb01380.x.

Booth, D.B., J.R. Karr, S. Schauman, C.P. Konrad, S.A. Morley, M.G. Larson, and S.J. Burges (2004). Reviving urban streams: land use, hydrology, biology, and human behavior. *Journal of the American Water Resources Association (JAWRA)* **40**(5):1351-1364, DOI: 10.1111/j.1752-1688.2004.tb01591.x.

Brooks, A.J., B.C. Chessman, and T. Haeusler (2011) Macroinvertebrate traits distinguish unregulated rivers subject to water abstraction. *Journal of the North American Benthological Society* **30**(2):419-435, DOI: 10.1899/10-074.1.

Bunn, S.E. and A.H. Arthington (2002). Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. *Environmental Management*, **30**(4):492–507, DOI: 10.1007/s00267-002-2737-0.

Cairns, J., and J.R. Pratt (1993). A History of Biological Monitoring Using Benthic Macroinvertebrates. *Freshwater Biomonitoring and Benthic Macroinvertebrates*. Ed. Rosenberg, D.M. and V.H. Resh. Chapman and Hall, New York, p.10-27.

California Department of Finance (2014). State and County Population Projections, 2010-2060. State of California Report P-1 (County), Sacramento, California, URL: http://www.dof.ca.gov/research/demographic/reports/projections/P-1/.

California Irrigation Management Information System (CIMIS) (2015). Station Reports. California Department of Water Resources, State of California, URL: http://www.cimis.water.ca.gov/.

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

Chow, V.T., D.R. Maidment, and L.W. Mays (1988). Applied Hydrology International Edition. McGraw-Hill Publishing Company, New York.

Datry, T. (2012). Benthic and Hyporheic invertebrate assemblages along a flow intermittence gradient: effects of duration of dry events. *Freshwater Biology* **57**(3):563-574, DOI: 10.1111/j.1365- 2427.2011.02725.x.

Eberhart, S.R. (2014). On the development of flow-ecology relationships for streams in coastal watersheds of southern California. M.S. Thesis, Colorado State University, Fort Collins, CO, 74p., URL: http://hdl.handle.net/10217/88519.

Falcone, J. (2011). GAGES-II: Geospatial Attributes of Gages for Evaluating Streamflow. U.S. Department of the Interior, U.S. Geological Survey, Reston, Virginia, URL: http://water.usgs.gov/GIS/metadata/usgswrd/XML/gagesII_Sept2011.xml.

Federal Emergency Management Agency (2013). Hydrologic Models Meeting the Minimum Requirement of National Flood Insurance Program. US Department of Homeland Security, URL: https://www.fema.gov/hydrologic-models-meeting-minimum-requirement-national-flood-insurance-program. Accessed July 2014.

Forman, R.T.T. (2008). The urban region: natural systems in our place, our nourishment, our home range, our future. *Landscape Ecology* **23**(3):251-253, DOI: 10.1007/s10980-008-9209-8.

Fry, J., Xian, G., Jin, S., Dewitz, J., Homer, C., Yang, L., Barnes, C., Herold, N., and Wickham, J. (2011). Completion of the 2006 National Land Cover Database for the Conterminous United States. *Photogrammetric Engineering & Remote Sensing (PE&RS)* **77**(9):858-864. URL: http://www.mrlc.gov/nlcd2006.php.

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**(1):51-81, DOI: 10.1146/annurev.ecolsys.30.1.51.

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):5-11.

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

Gupta, H.V., T. Wagener, and Y. Liu (2008). Reconciling theory with observations: elements of a diagnostic approach to model evaluation. *Hydrological Processes* **22**(18):3802-3813, DOI: 10.1002/hyp.6989.

Hawley, R.J., and B.P. Bledsoe (2011). How do flow peaks and durations change in suburbanizing semi-arid watersheds? A southern California case study. *Journal of Hydrology* **405**(1–2):69–82, DOI: 10.1016/j.jhydrol.2011.05.011.

Hawley, R.J., B.P. Bledsoe, E.D. Stein, and B.E. Haines (2012). Channel Evolution Model of Semiarid Stream Response to Urban-Induced Hydromodification. *Journal of the American Water Resources Association (JAWRA)*, **48**(4):722-744, DOI: 10.1111/j.1752-1688.2012.00645.x.

Helms, B.S., J.E. Schoonover, and J.W. Feminella (2009). Seasonal variability of landuse impacts on macroinvertebrate assemblages in streams of western Georgia, USA. *Journal of the North American Benthological Society*, **28**(4):991–1006, DOI: 10.1899/08-162.1.

Henrikson, J.A., J. Heasley, J.G. Kennen, and S. Nieswand (2006). Users' Manual for the Hydroecological Integrity Assessment Process Software (including the New Jersey Assessment Tools). Open File Report 2006-1093, U. S. Department of the Interior, U. S. Geological Survey, Reston, VA, 71 p., URL: https://www.fort.usgs.gov/sites/default/files/products/publications/21598/ 21598.pdf.

Hydrologic Engineering Center (2013). Hydrologic Modeling System HEC-HMS User's Manual Version 4.0. U.S. Army Corps of Engineers Report, 426p., URL: http://www.hec.usace.army.mil/software/hec-hms/documentation/HEC-HMS_Users_Manual_4.0.pdf.

Jacobson, R.B., S.R. Femmer, and R.A. McKenney (2001). Land-Use Changes and the Physical Habitat of Streams—A Review with Emphasis on Studies within the U.S. Geological Survey Federal-State Cooperative Program. U.S. Geological Survey Circular 1175, Reston, Virginia, 63 p., URL: http://pubs.usgs.gov/circ/circ1175/pdf/circ1175.pdf.

Jain, S.K., and K.P. Sudheer (2008). Fitting of Hydrologic Models: A Close Look at the Nash-Sutcliffe Index. *Journal of Hydrologic Engineering* **13**(10):981-986, DOI: 10.1061/(ASCE)1084-0699(2008)13:10(981).

Jakeman, A. J., and G. M. Hornberger (1993). How much complexity is warranted in a rainfall-runoff model? *Water Resources Research* **29**(8):2637-2649, DOI: 10.1029/93WR00877.

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

Kennard, M.J., S.J. MacKay, B.J. Pusey, J.D. Olden, and N. Marsh (2010). Quantifying Uncertainty in Estimation of Hydrologic Metrics for Ecohydrological Studies. *River Research and Applications* **26**(2):137-156, DOI: 10.1002/rra.1249.

Kennen, J.G., and M.L. Riskin (2010). Evaluating Effects of Potential Changes in Streamflow Regime on Fish and Aquatic-Invertebrate Assemblages in the New Jersey Pinelands. U.S. Geological Survey Scientific Investigations Report 2010-5079, 34 p., URL: http://pubs.usgs.gov/sir/2010/5079/.

Kennen, J.G., K. Riva-Murray, and K.M. Beaulieu (2010). Determining hydrologic factors that influence stream macroinvertebrate assemblages in the northeastern US. *Ecohydrology* **3**(1):88–106, DOI: 10.1002/eco.99.

Kennen, J.G., M.L. Riskin, P.A. Reilly, and S.J. Colarullo (2013). Method to Support Total Maximum Daily Load Development Using Hydrologic Alteration as a Surrogate to Address Aquatic-Life Impairment in New Jersey Streams. U.S. Geological Survey Scientific Investigations Report 2013-5089, 86p., URL: http://pubs.usgs.gov/sir/2013/5089/pdf/sir2013-5089.pdf.

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

Kolkwitz, R. and M. Marsson (1909). Ökologie der tierischen Saprobien. Beiträge zur Lehre von der biologischen Gewässerbeurteilung. *Internationale Revue der gesamten Hydrobiologie und Hydrographie* **2**(1-2):126-152, DOI: 10.1002/iroh.19090020108.

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

Konrad, C.P., D.B. Booth, and S.J. Burges (2005). Effects of urban development in Puget Lowland, Washington, on interannual streamflow patterns: consequences for channel form and streambed disturbance. *Water Resources Research* **41**(7):1-15, DOI: 10.1029/2005WR004097.

Konrad, C.P., A.M. Brasher, and J.T. May (2008). Assessing streamflow characteristics as limiting factors on benthic invertebrate assemblages in streams across the western United States. *Freshwater Biology*, **53**(10):1983–1998, DOI: 10.1111/j.1365-2427.2008.02024.x.

Knight, R.R., and T.F. Cuffney (2012). Invertebrate Response to Changes in Streamflow Hydraulics in Two Urban Areas in the United States. Scientific Investigations Report 2012-5035, U. S. Department of the Interior, U. S. Geological Survey, National Water-Quality Assessment Program, Reston, VA, 19 p., URL: http://pubs.usgs.gov/sir/2012/5035/pdf/2012-5035.pdf.

Liaw, A. and M. Wiener (2002). Classification and Regression by randomForest. *R News* **2**(3):18-22, URL: http://cran.r-project.org/web/packages/randomForest/index.html.

Likens, G.H. (1984). Beyond the shoreline: a watershed-ecosystem approach. *Verhandlungen der Internationalen Vereinigung fur Theoretische und Angewandte Limnologie* **22**:1-22.

Mallows, C.L. (1973). Some Comments on C_P. *Technometrics* **15**(4):661-675, DOI: 10.2307/1267380.

McElravy, E.P. (1988) Temporal variability in abundance of aquatic insects: a comparison of temperate and tropical habitats. Ph.D. Thesis, University of California, Berkeley, CA.

Murphy, J.C., R.R. Knight, W.J. Wolfe, and W.S. Gain (2013). Predicting Ecological Flow Regime at Ungaged Sites: A Comparison of Methods. *River Research and Applications* **29**(5):660-669, DOI: 10.1002/rra.2570.

Nash, J.E., and J.V. Sutcliffe (1970). River flow forecasting through conceptual models part I – A discussion of principles. *Journal of Hydrology* **10**(3):282-290, DOI: 10.1016/0022-1694(70)90255-6.

The Nature Conservancy (TNC) Conservation Gateway (2015). Ecological Limits of Hydrologic Alteration (ELOHA), URL: https://www.conservationgateway.org/ConservationPractices/Freshwater/EnvironmentalFlows/MethodsandTools/ELOHA/Pages/ecological-limits-hydrolo.aspx. Accessed May 2015.

NBS/Lowry Engineers & Planners, Inc. and George V. Sabol Consulting Engineers, Inc. (1993). Highway Drainage Design Manual Hydrology. Arizona Department of Transportation Report, 338p., URL: http://azdot.gov/docs/business/highway-drainage-design-manual-hydrology.

Ode, P.R., A.C. Rehn, and J.T. May (2005). A quantitative tool for assessing the integrity of southern coastal California streams. *Environmental Management*, **35**(4):493–504, DOI: 10.1007/s00267-004-0035-8.

Poff, N.L. and J.V. Ward (1989). Implications of streamflow variability and predictability for lotic community structure: a regional analysis of streamflow patterns. *Canadian Journal of Fisheries and Aquatic Sciences* **46**(10):1805-1818, DOI: 10.1139/f89-228.

Poff, N.L. (1996). A hydrogeography of unregulated streams in the United States and an examination of scale-dependence in some hydrological descriptors. *Freshwater Biology* **36**(1):71-91, DOI: 10.1046/j.1365-2427.1996.00073.x.

Poff, N.L., J.D. Allen, M.B. Bain, J.R. Karr, K.L. Prestegaard, B.D. Richter, R.E. Sparks, and J.C. Stromberg (1997). The Natural Flow Regime. *BioScience* **47**(11):769-784, DOI: 10.2307/1313099.

Poff, N.L., B.P. Bledsoe, and C.O. Cuhaciyan (2006a). Hydrologic variation with land use across the contiguous United States: Geomorphic and ecological consequences for stream ecosystems. *Geomorphology*, **79**(3-4):264-285, DOI: 10.1016/j.geomorph.2006.06.032.

Poff, N. L., J. D. Olden, N. K. Vieira, D. S. Finn, M. P. Simmons, and B. C. Kondratieff (2006b). Functional trait niches of North American lotic insects: traits-based ecological applications in light of phylogentic relationships. *Journal of the North American Benthological Society*, **25**(4):730–755, DOI: 10.1899/0887-3593(2006)025[0730:FTNONA]2.0.CO;2.

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

Poff, N.L., B.D. Richter, A.H. Arthington, S.E. Bunn, R.J. Naiman, E. Kendy, M. Acreman, C. Aspe, B.P. Bledsoe, M.C. Freeman, J. Henriksen, R.B. Jacobson, J.G. Kennen, D.M. Merritt, J.H. O'Keeffe, J.D. Olden, K. Rogers, R.E. Tharme, and A. Warner (2010). The ecological limits of hydrologic alteration (ELOHA): a new framework for developing regional environmental flow standards. *Freshwater Biology* **55**(1):147-170, DOI: 10.1111/j.1365-2427.2009.02204.x.

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):194–205, DOI: 10.1111/j.1365-2427.2009.02272.x.

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

Resh, V.H., A.V. Brown, A.P. Covich, M.E. Gurtz, H.W. Li, G.W. Minshall, S.R. Reice, A.L. Sheldon, J.B. Wallace, and R.C. Wissmar (1988). The role of disturbance in stream ecology. *Journal of the North American Benthological Society* **7**(4):433-455, DOI: 10.2307/1467300.

Resh, V.H., R.H. Norris, and M.T. Barbour (2006). Design and implementation of rapid assessment approaches for water resource monitoring using benthic macroinvertebrates. *Australian Journal of Ecology* **20**(1):108-121, DOI: 10.1111/j.1442-9993.1995.tb00525.x.

Richter, B.D., J.V. Baungartner, J. Powell, and D.P. Braun (1996). A method for assessing hydrologic alteration within ecosystems. *Conservation Biology* **10**(4):1163-1174, DOI: 10.1046/j.1523- 1739.1996.10041163.x.

Schaake, J.C. (1981). Summary of River Forecasting Raingage Network Density Requirements, unpublished, URL: http://www.nws.noaa.gov/oh/mopex/raingage%20density%20requirement.htm.

SAS Institute Inc. (2015). SAS University Edition Software. Cary, NC.

Soar, P.J. and C.R. Thorne (2001). Channel Restoration Design for Meandering Rivers. U.S. Army Corps of Engineers Report, 454p., URL: http://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/stelprdb1043218.pdf.

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

Tierra Grande International Inc., and G.V. Sabol (2008). Hydrologic Basin Response Parameter Estimation Guidelines. State of Colorado Office of the State Engineer Dam Safety Branch Report, 61p., URL: http://water.state.co.us/DWRIPub/Documents/HyBasinRsp200805.pdf.

U.S. Census Bureau (2012). Growth in Urban Population Outpaces Rest of Nation, Census Bureau Reports. Report CB12-50, URL: https://www.census.gov/newsroom/releases/archives/2010_census/cb12-50.html.

U.S. Department of Agriculture (USDA) (2015). Annual cash receipts by commodity. Economic Research Service Farm Income and Wealth Statistics, URL: http://www.ers.usda.gov/data-products/farm-income-and-wealth-statistics/annual-cash-receipts-by-commodity.aspx.

U.S. Environmental Protection Agency (EPA) (2014). Ecoregions of North America. EPA Western Ecology Division, URL: http://www.epa.gov/wed/pages/ecoregions/na_eco.htm. Accessed August 6 2014.

U. S. Geological Survey (USGS) (2014). Instantaneous Data Archive (IDA). U. S. Department of the Interior, U. S. Geological Survey, URL: http://ida.water.usgs.gov/ida/.

U. S. Geological Survey (USGS) (2015). National Water Information System (NWIS). U. S. Department of the Interior, U. S. Geological Survey, URL: http://waterdata.usgs.gov/nwis.

Vis, M., R. Knight, S. Pool, W. Wolfe, and J. Seibert (2015). Model Calibration Criteria for Estimating Ecological Flow Characteristics. *Water* **7**(5):2358-2381, DOI: 10.3390/w7052358.

Walsh, C.J., A.H. Roy, J.W. Feminella, P.D. Cottingham, P.M. Groffman, and R.P. Morgan (2005). The urban stream syndrome: current knowledge and the search for a cure. *Journal of the North American Benthological Society* **24**(3):706-723, DOI: 10.2307/4095691.

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

Zickovich, J.M and A.J. Bohonak (2007). Dispersal ability and genetic structure in aquatic invertebrates: a comparative study in southern California streams and reservoirs. *Freshwater Biology* **52**(10)1982-1996, DOI: 10.1111/j.1365-2427.2007.01822.x.

Appendix I: Flow metric accuracy relative to the three WY antecedent gage data and flow metric correlation matrices

Table A1.1: Accuracy of flow metrics from rainfall-runoff models and WY 2005-2007 gage data compared to the three WY antecedent gage data. Italicized average metric errors at the bottom indicate the calibration approach with the lowest average metric error.

Flow Metric	Mean % difference Gage WY 2005-2007	Median % difference Gage WY 2005-2007	Mean % error NSE Calibration	Median % error NSE Calibration	Mean % error RBI Calibration	Median % error RBI Calibration	Mean % error < 1 cfs Calibration	Median % error < 1 cfs Calibration	Mean % error RBI and < 1 cfs Combined Calibration	Median % error RBI and < 1 cfs Combined Calibration
RBI_1hr	24.1	22.4	7.9	3.2	27.4	25.0	15.4	52.3	27.7	19.4
cfs_1hr	1.0	23.9	25.0	146.8	127.8	361.0	0.1	27.9	0.8	26.9
Dec	34.1	98.5	48.7	80.0	95.5	100.0	128.9	73.8	371.8	48.3
Jan	25.7	84.2	56.6	81.5	84.9	100.0	165.3	35.2	554.5	64.5
Feb	80.5	110.0	74.1	86.6	84.7	100.0	57.7	63.3	291.7	58.4
SevenDayMax	61.2	99.6	67.4	96.5	13.5	94.2	10.7	86.6	296.7	211.3
NumZeroDays	59.8	100	26.9	100	231.0	100	69.0	100	17.3	100
DateMax	30.0	6.1	14.0	20.8	16.5	12.5	5.1	20.8	14.0	4.2
LowPulseCount	11.5	0.0	84.5	100.0	100.0	100.0	68.9	75.0	35.8	25.0
HighPulseCount	10.1	18.2	2.2	20.0	69.8	140.0	18.0	20.0	106.7	180.0
RiseRate	52.8	40.0	5818.6	1236.3	5463.0	7290.0	3228.4	1072.3	18090.8	25520.0
FallRate	46.7	50.0	232.0	15.5	883.6	586.3	92.5	82.8	219.7	271.3
NumReversals	14.3	29.5	68.1	76.9	63.5	67.9	57.6	65.4	30.1	23.1
ExtremeLowDuration	27.3	0.0	9.3	166.7	96.2	6.7	22.1	213.3	72.7	13.3
ExtremeLowTiming	11.2	9.0	21.3	45.3	45.0	71.8	6.0	19.2	4.1	7.7
ExtremeLowFreq	14.6	40.0	22.9	66.7	141.0	100.0	18.9	66.7	27.6	0.0
HighFlowDuration	58.5	40.0	64.6	33.3	62.9	33.3	17.0	50.0	43.6	33.3
LargeFloodDuration	66.1	71.9	64.0	25.7	28.6	64.9	134.7	216.2	39.9	64.9
SmallFloodPeak	23.3	42.6	35.0	70.2	100.5	33.6	37.1	78.1	435.7	400.5
SmallFloodRiseRate	35.0	55.4	6.9	49.6	163.7	146.7	201.1	25.4	496.0	1006.5
SmallFloodFallRate	55.8	28.4	188.6	2.5	265.1	419.3	333.4	3.2	440.7	1029.4

LargeFloodRiseRate	99.0	106.9	21.0	33.0	19.6	9.3	57.7	73.0	290.0	100.0
All Low Flow	20.9	16.5	31.6	100.0	123.5	100.0	30.8	70.8	26.4	*19.2*
All Flashiness	46.8	40.0	906.2	33.0	983.7	146.7	569.4	*65.4*	2799.3	271.3
All	38.3	40.0	316.3	68.4	372.0	100.0	215.7	*66.0*	995.8	61.4

Table A1.2: Three WY antecedent gage data flow metric correlation matrix with correlations statistically significant at the p = 0.1 level denoted in italics.

Flow Metric	RBI_1hr	cfs_1hr	Dec	Jan	Feb	7DayMax	#ZeroDays	DateMax	LowPulseCount	HighPulseCount	Rise Rate	Fall Rate	#Reversals	Ext Low Dur	Ext Low Time	Ext Low Freq	High Dur	SmallFloodPeak	SmallFloodRise	SmallFloodFall	LargeFloodDur	LargeFloodRise
RBI_1hr	-	-0.12	0.32	0.28	0.00	0.29	-0.09	-0.07	*0.37*	*0.64*	0.06	-0.19	*0.45*	*-0.60*	0.03	*0.54*	-0.31	0.26	*0.38*	*-0.47*	-0.25	-0.08
cfs_1hr	-0.12	-	*-0.44*	*-0.48*	*-0.49*	*-0.56*	*0.87*	*-0.39*	*-0.66*	*-0.59*	*-0.37*	*0.50*	*-0.75*	*0.53*	*-0.69*	*-0.44*	*0.49*	*-0.56*	*-0.58*	*0.36*	*0.56*	*-0.45*
Dec	0.32	*-0.44*	-	*0.96*	*0.80*	*0.70*	*-0.38*	-0.22	*0.52*	*0.58*	*0.62*	*-0.85*	*0.47*	-0.22	0.06	*0.35*	-0.14	*0.64*	*0.42*	*-0.80*	-0.11	0.20
Jan	0.28	*-0.48*	*0.96*	-	*0.87*	*0.78*	*-0.38*	-0.04	*0.47*	*0.57*	*0.79*	*-0.95*	*0.45*	-0.16	0.10	0.29	-0.24	*0.74*	*0.48*	*-0.85*	-0.09	*0.37*
Feb	0.00	*-0.49*	*0.80*	*0.87*	-	*0.75*	-0.33	-0.01	0.26	0.26	*0.83*	*-0.92*	0.21	0.03	0.17	0.06	-0.10	*0.75*	*0.41*	*-0.57*	-0.03	*0.55*
7Day Max	0.29	*-0.56*	*0.70*	*0.78*	*0.75*	-	*-0.44*	0.19	*0.65*	*0.53*	*0.76*	*-0.80*	*0.47*	-0.16	0.32	*0.44*	-0.28	*0.97*	*0.81*	*-0.53*	-0.33	*0.75*
#Zero Days	-0.09	*0.87*	*-0.38*	*-0.38*	-0.33	*-0.44*	-	-0.24	*-0.69*	*-0.59*	-0.16	0.33	*-0.84*	*0.59*	*-0.77*	*-0.41*	*0.68*	*-0.44*	*-0.45*	0.31	*0.56*	-0.32
Date Max	-0.07	*-0.39*	-0.22	-0.04	-0.01	0.19	-0.24	-	0.15	0.13	0.28	-0.08	0.23	0.01	0.28	0.15	*-0.39*	0.22	*0.47*	0.06	-0.12	*0.44*
Low Pulse Count	*0.37*	*-0.66*	*0.52*	*0.47*	0.26	*0.65*	*-0.69*	0.15	-	*0.73*	0.28	*-0.40*	*0.81*	*-0.60*	*0.40*	*0.78*	*-0.49*	*0.56*	*0.63*	*-0.37*	*-0.57*	0.31
High Pulse Count	*0.64*	*-0.59*	*0.58*	*0.57*	0.26	*0.53*	*-0.59*	0.13	*0.73*	-	0.31	*-0.47*	*0.86*	*-0.60*	*0.34*	*0.79*	*-0.56*	*0.47*	*0.51*	*-0.68*	*-0.54*	0.09
Rise Rate	0.06	*-0.37*	*0.62*	*0.79*	*0.83*	*0.76*	-0.16	0.28	0.28	0.31	-	*-0.92*	0.16	0.14	0.10	0.15	-0.25	*0.73*	*0.50*	*-0.53*	0.02	*0.62*
Fall Rate	-0.19	*0.50*	*-0.85*	*-0.95*	*-0.92*	*-0.80*	0.33	-0.08	*-0.40*	*-0.47*	*-0.92*	-	*-0.36*	0.06	-0.13	-0.20	0.26	*-0.78*	*-0.50*	*0.73*	0.04	*-0.52*
#Reversals	*0.45*	*-0.75*	*0.47*	*0.45*	0.21	*0.47*	*-0.84*	0.23	*0.81*	*0.86*	0.16	*-0.36*	-	*-0.69*	*0.54*	*0.64*	*-0.67*	*0.45*	*0.49*	*-0.53*	*-0.61*	0.17
ExtLowDur	*-0.60*	*0.53*	-0.22	-0.16	0.03	-0.16	*0.59*	0.01	*-0.60*	*-0.60*	0.14	0.06	*-0.69*	-	*-0.39*	*-0.50*	*0.55*	-0.17	-0.19	0.30	*0.45*	-0.01
ExtLow Time	0.03	*-0.69*	0.06	0.10	0.17	0.32	*-0.77*	0.28	*0.40*	*0.34*	0.10	-0.13	*0.54*	*-0.39*	-	0.26	*-0.64*	*0.34*	*0.39*	-0.01	*-0.56*	*0.38*

ExtLow Freq	*0.54*	*-0.44*	*0.35*	0.29	0.06	*0.44*	*-0.41*	0.15	*0.78*	*0.79*	0.15	-0.20	*0.64*	*-0.50*	0.26	-	*-0.39*	0.31	*0.52*	-0.31	*-0.56*	0.01
High Dur	-0.31	*0.49*	-0.14	-0.24	-0.10	-0.28	*0.68*	*-0.39*	*-0.49*	*-0.56*	-0.25	0.26	*-0.67*	*0.55*	*-0.64*	*-0.39*	-	-0.28	-0.32	0.30	0.33	-0.24
Small Flood Peak	0.26	*-0.56*	*0.64*	*0.74*	*0.75*	*0.97*	*-0.44*	0.22	*0.56*	*0.47*	*0.73*	*-0.78*	*0.45*	-0.17	*0.34*	0.31	-0.28	-	*0.79*	*-0.54*	-0.30	*0.82*
Small Flood Rise	*0.38*	*-0.58*	*0.42*	*0.48*	*0.41*	*0.81*	*-0.45*	*0.47*	*0.63*	*0.51*	*0.50*	*-0.50*	*0.49*	-0.19	*0.39*	*0.52*	-0.32	*0.79*	-	*-0.38*	*-0.38*	*0.64*
Small Flood Fall	*-0.47*	*0.36*	*-0.80*	*-0.85*	*-0.57*	*-0.53*	0.31	0.06	*-0.37*	*-0.68*	*-0.53*	*0.73*	*-0.53*	0.30	-0.01	-0.31	0.30	*-0.54*	*-0.38*	-	0.14	-0.09
Large Flood Dur	-0.25	*0.56*	-0.11	-0.09	-0.03	-0.33	*0.56*	-0.12	*-0.57*	*-0.54*	0.02	0.04	*-0.61*	*0.45*	*-0.56*	*-0.56*	0.33	-0.30	*-0.38*	0.14	-	-0.20
Large Flood Rise	-0.08	*-0.45*	0.20	*0.37*	*0.55*	*0.75*	-0.32	*0.44*	0.31	0.09	*0.62*	*-0.52*	0.17	-0.01	*0.38*	0.01	-0.24	*0.82*	*0.64*	-0.09	-0.20	-

Table A1.3: WY 2005-2007 gage data flow metric correlation matrix with correlations statistically significant at the p = 0.1 level denoted in italics.

Flow Metric	RBI_1hr	cfs_1hr	Dec	Jan	Feb	7DayMax	#ZeroDays	DateMax	LowPulseCount	HighPulseCount	Rise Rate	Fall Rate	#Reversals	Ext Low Dur	Ext Low Time	Ext Low Freq	High Dur	SmallFloodPeak	SmallFloodRise	SmallFloodFall	LargeFloodDur	LargeFloodRise
RBI_1hr	-	-0.14	*0.34*	0.27	*0.37*	0.30	0.08	0.22	*0.62*	*0.71*	*0.42*	*-0.47*	*0.56*	0.07	-0.07	*0.63*	0.08	*0.44*	0.23	*-0.61*	*-0.34*	-0.19
cfs_1hr	-0.14	-	*-0.45*	*-0.54*	*-0.48*	-0.35	*0.87*	-0.27	*-0.57*	*-0.50*	*-0.56*	*0.50*	*-0.68*	*0.76*	*-0.62*	*-0.46*	-0.16	*-0.44*	*-0.35*	*0.46*	*0.80*	*-0.65*
Dec	*0.34*	*-0.45*	-	*0.98*	*1.00*	*0.46*	-0.34	-0.26	*0.65*	*0.61*	*0.96*	*-0.98*	*0.54*	-0.30	0.21	0.11	-0.22	*0.50*	0.08	*-0.75*	*-0.53*	*0.59*
Jan	0.27	*-0.54*	*0.98*	-	*0.99*	*0.44*	*-0.40*	-0.22	*0.62*	*0.58*	*0.97*	*-0.96*	*0.52*	*-0.36*	0.28	0.12	-0.22	*0.49*	0.09	*-0.72*	*-0.58*	*0.71*
Feb	*0.37*	*-0.48*	*1.00*	*0.99*	-	*0.46*	*-0.36*	-0.20	*0.69*	*0.65*	*0.98*	*-0.99*	*0.57*	-0.32	0.26	0.18	-0.21	*0.52*	0.10	*-0.77*	*-0.58*	*0.60*
7Day Max	0.30	*-0.35*	*0.46*	*0.44*	*0.46*	-	-0.27	-0.08	*0.34*	0.22	*0.53*	*-0.51*	0.22	-0.21	0.20	-0.03	0.15	*0.97*	*0.89*	*-0.82*	-0.26	*0.43*
#Zero Days	0.08	*0.87*	-0.34	*-0.40*	*-0.36*	-0.27	-	-0.16	*-0.54*	*-0.46*	*-0.39*	*0.36*	*-0.70*	*0.94*	*-0.79*	*-0.36*	0.06	*-0.34*	-0.29	*0.35*	*0.80*	*-0.51*
Date Max	0.22	-0.27	-0.26	-0.22	-0.20	-0.08	-0.16	-	0.31	0.32	-0.16	0.17	0.27	-0.14	*0.35*	*0.74*	0.30	0.01	0.10	-0.02	*-0.40*	-0.14
Low Pulse Count	*0.62*	*-0.57*	*0.65*	*0.62*	*0.69*	*0.34*	*-0.54*	0.31	-	*0.92*	*0.70*	*-0.73*	*0.91*	*-0.46*	*0.40*	*0.73*	-0.04	*0.50*	0.18	*-0.75*	*-0.79*	0.15
High Pulse	*0.71*	*-0.50*	*0.61*	*0.58*	*0.65*	0.22	*-0.46*	0.32	*0.92*	-	*0.62*	*-0.67*	*0.89*	*-0.44*	*0.44*	*0.73*	-0.21	*0.37*	0.06	*-0.66*	*-0.79*	0.14

Count																						
Rise Rate	*0.42*	*-0.56*	*0.96*	*0.97*	*0.98*	*0.53*	*-0.39*	-0.16	*0.70*	*0.62*	-	*-0.99*	*0.59*	-0.33	0.25	0.24	-0.11	*0.61*	0.23	*-0.83*	*-0.62*	*0.64*
Fall Rate	*-0.47*	*0.50*	*-0.98*	*-0.96*	*-0.99*	*-0.51*	*0.36*	0.17	*-0.73*	*-0.67*	*-0.99*	-	*-0.62*	0.33	-0.25	-0.26	0.16	*-0.59*	-0.18	*0.83*	*0.62*	*-0.55*
#Reversals	*0.56*	*-0.68*	*0.54*	*0.52*	*0.57*	0.22	*-0.70*	0.27	*0.91*	*0.89*	*0.59*	*-0.62*	-	*-0.67*	*0.52*	*0.76*	-0.11	*0.38*	0.14	*-0.60*	*-0.87*	0.15
ExtLowDur	0.07	*0.76*	-0.30	*-0.36*	-0.32	-0.21	*0.94*	-0.14	*-0.46*	*-0.44*	-0.33	0.33	*-0.67*	-	*-0.88*	*-0.34*	0.23	-0.28	-0.22	0.30	*0.79*	*-0.46*
ExtLow Time	-0.07	*-0.62*	0.21	0.28	0.26	0.20	*-0.79*	*0.35*	*0.40*	*0.44*	0.25	-0.25	*0.52*	*-0.88*	-	*0.36*	-0.27	0.26	0.21	-0.28	*-0.78*	*0.42*
ExtLow Freq	*0.63*	*-0.46*	0.11	0.12	0.18	-0.03	*-0.36*	*0.74*	*0.73*	*0.73*	0.24	-0.26	*0.76*	*-0.34*	*0.36*	-	0.15	0.17	0.07	*-0.35*	*-0.70*	-0.15
High Dur	0.08	-0.16	-0.22	-0.22	-0.21	0.15	0.06	0.30	-0.04	-0.21	-0.11	0.16	-0.11	0.23	-0.27	0.15	-	0.16	0.29	-0.07	0.13	-0.17
Small Flood Peak	*0.44*	*-0.44*	*0.50*	*0.49*	*0.52*	*0.97*	*-0.34*	0.01	*0.50*	*0.37*	*0.61*	*-0.59*	*0.38*	-0.28	0.26	0.17	0.16	-	*0.89*	*-0.90*	*-0.41*	*0.40*
Small Flood Rise	0.23	*-0.35*	0.08	0.09	0.10	*0.89*	-0.29	0.10	0.18	0.06	0.23	-0.18	0.14	-0.22	0.21	0.07	0.29	*0.89*	-	*-0.62*	-0.21	0.29
Small Flood Fall	*-0.61*	*0.46*	*-0.75*	*-0.72*	*-0.77*	*-0.82*	*0.35*	-0.02	*-0.75*	*-0.66*	*-0.83*	*0.83*	*-0.60*	0.30	-0.28	*-0.35*	-0.07	*-0.90*	*-0.62*	-	*0.58*	*-0.37*
Large Flood Dur	*-0.34*	*0.80*	*-0.53*	*-0.58*	*-0.58*	-0.26	*0.80*	*-0.40*	*-0.79*	*-0.79*	*-0.62*	*0.62*	*-0.87*	*0.79*	*-0.78*	*-0.70*	0.13	*-0.41*	-0.21	*0.58*	-	*-0.44*
Large Flood Rise	-0.19	*-0.65*	*0.59*	*0.71*	*0.60*	*0.43*	*-0.51*	-0.14	0.15	0.14	*0.64*	*-0.55*	0.15	*-0.46*	*0.42*	-0.15	-0.17	*0.40*	0.29	*-0.37*	*-0.44*	-

Table A1.4: NSE calibration flow metric correlation matrix with correlations statistically significant at the p = 0.1 level denoted in italics.

Flow Metric	**RBI_1hr**	**cfs_1hr**	**Dec**	***Jan***	**Feb**	**7DayMax**	**#ZeroDays**	**DateMax**	**LowPulseCount**	**HighPulseCount**	**Rise Rate**	**Fall Rate**	**#Reversals**	**Ext Low Dur**	**Ext Low Time**	**Ext Low Freq**	**High Dur**	**SmallFloodPeak**	**SmallFloodRise**	**SmallFloodFall**	**LargeFloodDur**	**LargeFloodRise**
RBI_1hr	-	-0.02	0.23	*0.27*	0.25	0.30	-0.19	*0.45*	*0.35*	-0.08	*0.54*	0.00	0.02	-0.08	0.03	*-0.57*	-0.16	0.27	0.14	-0.29	*-0.36*	0.09
cfs_1hr	-0.02	-	-0.29	*-0.32*	*-0.38*	*-0.43*	*0.78*	-0.07	-0.25	*-0.61*	0.01	*0.43*	*-0.41*	*0.37*	-0.32	*-0.46*	-0.31	*-0.48*	*-0.53*	*0.47*	*0.49*	-0.33
Dec	0.23	-0.29	-	*0.96*	*0.95*	*0.75*	-0.20	*0.64*	*0.95*	*0.34*	-0.15	0.18	*0.53*	0.10	0.27	-0.32	-0.16	*0.79*	*0.43*	*-0.49*	-0.17	0.14
Jan	0.27	-0.32	*0.96*	-	*0.98*	*0.88*	-0.22	*0.68*	*0.98*	0.28	-0.16	0.18	*0.64*	0.06	0.28	*-0.34*	-0.10	*0.87*	*0.58*	*-0.64*	-0.07	*0.37*

Feb	0.25	*-0.38*	*0.95*	*0.98*	-	*0.90*	-0.27	*0.64*	*0.97*	*0.43*	-0.16	0.17	*0.74*	-0.02	*0.40*	-0.28	-0.06	*0.89*	*0.63*	*-0.66*	-0.07	*0.44*
7Day Max	0.30	*-0.43*	*0.75*	*0.88*	*0.90*	-	*-0.35*	*0.57*	*0.87*	*0.39*	-0.10	-0.06	*0.75*	-0.19	0.27	-0.19	0.22	*0.98*	*0.88*	*-0.92*	-0.04	*0.72*
#Zero Days	-0.19	*0.78*	-0.20	*-0.22*	-0.27	*-0.35*	-	-0.32	-0.21	*-0.40*	-0.18	0.26	-0.24	0.13	-0.31	-0.21	-0.03	*-0.35*	*-0.41*	*0.40*	*0.38*	-0.26
Date Max	*0.45*	-0.07	*0.64*	*0.68*	*0.64*	*0.57*	-0.32	-	*0.73*	-0.20	*0.35*	*0.39*	0.13	*0.42*	0.02	*-0.65*	*-0.46*	*0.52*	0.25	*-0.39*	-0.26	0.12
Low Pulse Count	*0.35*	-0.25	*0.95*	*0.98*	*0.97*	*0.87*	-0.21	*0.73*	-	0.28	-0.10	0.26	*0.62*	0.02	0.23	*-0.40*	-0.12	*0.85*	*0.54*	*-0.63*	-0.10	*0.35*
High Pulse Count	-0.08	*-0.61*	*0.34*	*0.28*	*0.43*	*0.39*	*-0.40*	-0.20	0.28	-	-0.27	-0.22	*0.66*	*-0.62*	*0.52*	*0.50*	*0.45*	*0.44*	*0.45*	*-0.36*	-0.18	0.33
Rise Rate	*0.54*	0.01	-0.15	*-0.16*	-0.16	-0.10	-0.18	*0.35*	-0.10	-0.27	-	0.10	*-0.34*	0.30	-0.08	-0.27	-0.19	-0.13	-0.11	0.02	*-0.43*	-0.11
Fall Rate	0.00	*0.43*	0.18	*0.18*	0.17	-0.06	0.26	*0.39*	0.26	-0.22	0.10	-	0.12	0.28	0.03	*-0.40*	*-0.60*	-0.19	*-0.40*	*0.37*	0.26	-0.11
#Reversals	0.02	*-0.41*	*0.53*	*0.64*	*0.74*	*0.75*	-0.24	0.13	*0.62*	*0.66*	*-0.34*	0.12	-	*-0.38*	*0.62*	0.07	0.20	*0.71*	*0.70*	*-0.61*	0.26	*0.76*
ExtLowDur	-0.08	*0.37*	0.10	*0.06*	-0.02	-0.19	0.13	*0.42*	0.02	*-0.62*	0.30	0.28	*-0.38*	-	0.08	*-0.48*	*-0.70*	-0.19	*-0.34*	0.31	0.13	*-0.38*
ExtLow Time	0.03	-0.32	0.27	*0.28*	*0.40*	0.27	-0.31	0.02	0.23	*0.52*	-0.08	0.03	*0.62*	0.08	-	-0.07	-0.26	0.27	0.28	-0.14	0.26	*0.35*
ExtLow Freq	*-0.57*	*-0.46*	-0.32	*-0.34*	-0.28	-0.19	-0.21	*-0.65*	*-0.40*	*0.50*	-0.27	*-0.40*	0.07	*-0.48*	-0.07	-	*0.69*	-0.14	0.09	0.03	-0.16	0.05
High Dur	-0.16	-0.31	-0.16	*-0.10*	-0.06	0.22	-0.03	*-0.46*	-0.12	*0.45*	-0.19	*-0.60*	0.20	*-0.70*	-0.26	*0.69*	-	0.27	*0.50*	*-0.48*	-0.14	*0.37*
Small Flood Peak	0.27	*-0.48*	*0.79*	*0.87*	*0.89*	*0.98*	*-0.35*	*0.52*	*0.85*	*0.44*	-0.13	-0.19	*0.71*	-0.19	0.27	-0.14	0.27	-	*0.89*	*-0.92*	-0.11	*0.64*
Small Flood Rise	0.14	*-0.53*	*0.43*	*0.58*	*0.63*	*0.88*	*-0.41*	0.25	*0.54*	*0.45*	-0.11	*-0.40*	*0.70*	*-0.34*	0.28	0.09	*0.50*	*0.89*	-	*-0.97*	0.01	*0.86*
Small Flood Fall	-0.29	*0.47*	*-0.49*	*-0.64*	*-0.66*	*-0.92*	*0.40*	*-0.39*	*-0.63*	*-0.36*	0.02	*0.37*	*-0.61*	0.31	-0.14	0.03	*-0.48*	*-0.92*	*-0.97*	-	0.09	*-0.78*
Large Flood Dur	*-0.36*	*0.49*	-0.17	*-0.07*	-0.07	-0.04	*0.38*	-0.26	-0.10	-0.18	*-0.43*	0.26	0.26	0.13	0.26	-0.16	-0.14	-0.11	0.01	0.09	-	0.30
Large Flood Rise	0.09	-0.33	0.14	*0.37*	*0.44*	*0.72*	-0.26	0.12	*0.35*	0.33	-0.11	-0.11	*0.76*	*-0.38*	*0.35*	0.05	*0.37*	*0.64*	*0.86*	*-0.78*	0.30	-

Table A1.5: RBI calibration flow metric correlation matrix with correlations statistically significant at the p = 0.1 level denoted in italics.

Flow Metric	RBI_1hr	cfs_1hr	Dec	Jan	Feb	7DayMax	#ZeroDays	DateMax	LowPulseCount	HighPulseCount	Rise Rate	Fall Rate	#Reversals	Ext Low Dur	Ext Low Time	Ext Low Freq	High Dur	SmallFloodPeak	SmallFloodRise	SmallFloodFall	Large Flood Dur	LargeFloodRise
RBI_1hr	-	-0.18	0.33	-0.09	-0.28	0.00	-0.07	*0.37*	-	*0.35*	*0.44*	-0.24	-0.01	-0.19	-0.33	*0.43*	*0.49*	0.11	0.01	-0.27	-0.12	*-0.39*
cfs_1hr	-0.18	-	*-0.77*	0.00	-0.03	0.03	*0.93*	0.19	-	0.12	*-0.66*	*0.60*	0.05	0.12	-0.21	0.29	0.02	0.01	-0.11	*0.49*	-0.07	*-0.54*
Dec	0.33	*-0.77*	-	*0.50*	*0.36*	-0.10	*-0.57*	-0.06	-	*-0.35*	*0.61*	*-0.65*	*-0.39*	-0.03	0.01	-0.27	0.14	-0.10	-0.02	-0.27	-0.11	0.04
Jan	-0.09	0.00	*0.50*	-	*0.94*	-0.06	0.06	-0.01	-	-0.34	0.06	-0.21	-0.25	0.01	*0.34*	-0.20	0.05	-0.11	-0.09	0.12	-0.06	-0.16
Feb	-0.28	-0.03	*0.36*	*0.94*	-	-0.08	-0.07	-0.06	-	-0.29	0.05	-0.12	-0.10	-0.02	*0.59*	-0.28	-0.07	-0.13	-0.10	0.11	0.00	0.09
7Day Max	0.00	0.03	-0.10	-0.06	-0.08	-	0.01	-0.02	-	*0.46*	0.06	-0.24	*0.55*	-0.11	-0.01	*0.44*	0.18	*0.98*	*0.97*	*-0.71*	-0.18	0.13
#Zero Days	-0.07	*0.93*	*-0.57*	0.06	-0.07	0.01	-	0.22	-	0.05	*-0.62*	*0.36*	-0.10	0.17	*-0.48*	*0.35*	0.09	0.01	-0.10	*0.44*	-0.19	*-0.64*
Date Max	*0.37*	0.19	-0.06	-0.01	-0.06	-0.02	0.22	-	-	0.05	*0.45*	-0.03	-0.11	*0.36*	-0.11	0.15	-0.14	0.07	-0.05	-0.16	-0.33	*-0.38*
Low Pulse Count	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
High Pulse Count	*0.35*	0.12	*-0.35*	-0.34	-0.29	*0.46*	0.05	0.05	-	-	-0.11	-0.16	*0.90*	*-0.41*	0.00	*0.89*	*0.39*	*0.57*	*0.40*	*-0.56*	-0.12	0.31
Rise Rate	*0.44*	*-0.66*	*0.61*	0.06	0.05	0.06	*-0.62*	*0.45*	-	-0.11	-	*-0.46*	-0.20	0.06	0.07	-0.16	0.15	0.10	0.14	*-0.50*	*-0.35*	0.10
Fall Rate	-0.24	*0.60*	*-0.65*	-0.21	-0.12	-0.24	*0.36*	-0.03	-	-0.16	*-0.46*	-	-0.09	0.17	0.26	-0.27	-0.26	-0.32	*-0.36*	*0.71*	0.32	*-0.34*
#Reversals	-0.01	0.05	*-0.39*	-0.25	-0.10	*0.55*	-0.10	-0.11	-	*0.90*	-0.20	-0.09	-	*-0.35*	0.32	*0.68*	0.15	*0.62*	*0.49*	*-0.58*	0.03	*0.57*
ExtLowDur	-0.19	0.12	-0.03	0.01	-0.02	-0.11	0.17	*0.36*	-	*-0.41*	0.06	0.17	*-0.35*	-	-0.01	-0.32	*-0.51*	-0.14	-0.13	0.21	0.06	-0.24
ExtLow Time	-0.33	-0.21	0.01	*0.34*	*0.59*	-0.01	*-0.48*	-0.11	-	0.00	0.07	0.26	0.32	-0.01	-	*-0.35*	*-0.36*	-0.05	-0.03	-0.02	*0.36*	*0.50*
ExtLow Freq	*0.43*	0.29	-0.27	-0.20	-0.28	*0.44*	*0.35*	0.15	-	*0.89*	-0.16	-0.27	*0.68*	-0.32	*-0.35*	-	*0.53*	*0.56*	*0.37*	*-0.47*	*-0.35*	0.00
High Dur	*0.49*	0.02	0.14	0.05	-0.07	0.18	0.09	-0.14	-	*0.39*	0.15	-0.26	0.15	*-0.51*	*-0.36*	*0.53*	-	0.21	0.15	-0.22	*-0.50*	-0.16
Small Flood Peak	0.11	0.01	-0.10	-0.11	-0.13	*0.98*	0.01	0.07	-	*0.57*	0.10	-0.32	*0.62*	-0.14	-0.05	*0.56*	0.21	-	*0.95*	*-0.79*	-0.20	0.15
Small Flood	0.01	-0.11	-0.02	-0.09	-0.10	*0.97*	-0.10	-0.05	-	*0.40*	0.14	*-0.36*	*0.49*	-0.13	-0.03	*0.37*	0.15	*0.95*	-	*-0.77*	-0.14	0.23

Rise																						
Small Flood Fall	-0.27	*0.49*	-0.27	0.12	0.11	*-0.71*	*0.44*	-0.16	-	*-0.56*	*-0.50*	*0.71*	*-0.58*	0.21	-0.02	*-0.47*	-0.22	*-0.79*	*-0.77*	-	0.25	*-0.47*
Large Flood Dur	-0.12	-0.07	-0.11	-0.06	0.00	-0.18	-0.19	-0.33	-	-0.12	*-0.35*	0.32	0.03	0.06	*0.36*	*-0.35*	*-0.50*	-0.20	-0.14	0.25	-	0.11
Large Flood Rise	*-0.39*	*-0.54*	0.04	-0.16	0.09	0.13	*-0.64*	*-0.38*	-	0.31	0.10	*-0.34*	*0.57*	-0.24	*0.50*	0.00	-0.16	0.15	0.23	*-0.47*	0.11	-

Table A1.6: < 1 cfs Calibration Flow Metric Correlation Matrix with correlations statistically significant at the p = 0.1 level denoted in italics.

Flow Metric	RBI_1hr	cfs_1hr	Dec	Jan	Feb	7DayMax	#ZeroDays	DateMax	LowPulseCount	HighPulseCount	Rise Rate	Fall Rate	#Reversals	Ext Low Dur	Ext Low Time	Ext Low Freq	High Dur	SmallFloodPeak	SmallFloodRise	SmallFloodFall	Large Flood Dur	LargeFloodRise
RBI_1hr	-	-0.11	-0.23	-0.25	-0.24	-0.13	-0.19	0.27	0.06	0.14	*0.45*	0.07	-0.01	0.19	-0.02	-0.17	-0.03	-0.09	-0.18	*-0.48*	-0.34	0.03
cfs_1hr	-0.11	-	*-0.44*	*-0.35*	-0.33	*-0.37*	*0.74*	0.03	*-0.46*	0.18	-0.26	*0.39*	0.08	0.04	-0.11	*0.73*	*-0.40*	*-0.66*	*-0.51*	*0.53*	*0.36*	*-0.56*
Dec	-0.23	*-0.44*	-	*0.98*	*0.97*	*0.94*	-0.26	-0.22	*0.67*	0.09	*0.38*	-0.15	0.29	-0.18	0.31	-0.30	*0.75*	*0.78*	*0.43*	-0.07	*-0.35*	*0.36*
Jan	-0.25	*-0.35*	*0.98*	-	*1.00*	*0.97*	-0.21	-0.25	*0.62*	0.05	*0.42*	-0.21	0.22	-0.18	0.27	-0.24	*0.80*	*0.66*	0.29	0.03	*-0.37*	0.24
Feb	-0.24	-0.33	*0.97*	*1.00*	-	*0.98*	-0.20	-0.24	*0.61*	0.04	*0.42*	-0.22	0.20	-0.17	0.26	-0.22	*0.80*	*0.63*	0.25	0.05	*-0.38*	0.21
7Day Max	-0.13	*-0.37*	*0.94*	*0.97*	*0.98*	-	-0.23	-0.17	*0.63*	0.07	*0.45*	-0.27	0.24	-0.17	0.20	-0.24	*0.84*	*0.62*	0.18	-0.08	*-0.46*	0.19
#Zero Days	-0.19	*0.74*	-0.26	-0.21	-0.20	-0.23	-	-0.21	-0.33	*0.65*	-0.30	0.22	*0.34*	-0.28	*-0.37*	*0.96*	-0.12	*-0.42*	-0.32	*0.41*	0.11	*-0.42*
Date Max	0.27	0.03	-0.22	-0.25	-0.24	-0.17	-0.21	-	*0.40*	-0.03	0.01	0.23	-0.10	*0.36*	0.02	-0.29	*-0.37*	-0.24	*-0.37*	-0.08	-0.13	*-0.39*
Low Pulse Count	0.06	*-0.46*	*0.67*	*0.62*	*0.61*	*0.63*	-0.33	*0.40*	-	0.24	0.17	0.05	0.31	0.03	*0.36*	*-0.41*	*0.35*	*0.47*	0.08	-0.03	*-0.48*	0.02
High Pulse Count	0.14	0.18	0.09	0.05	0.04	0.07	*0.65*	-0.03	0.24	-	-0.09	-0.04	*0.56*	-0.32	*-0.38*	*0.64*	0.22	0.09	-0.04	-0.04	*-0.40*	0.02
Rise Rate	*0.45*	-0.26	*0.38*	*0.42*	*0.42*	*0.45*	-0.30	0.01	0.17	-0.09	-	-0.32	-0.28	0.00	-0.21	-0.26	*0.58*	0.31	0.14	*-0.39*	*-0.54*	0.29
Fall Rate	0.07	*0.39*	-0.15	-0.21	-0.22	-0.27	0.22	0.23	0.05	-0.04	-0.32	-	*0.34*	0.11	*0.61*	0.00	*-0.63*	-0.11	0.05	0.17	*0.54*	-0.32
#Reversals	-0.01	0.08	0.29	0.22	0.20	0.24	*0.34*	-0.10	0.31	*0.56*	-0.28	*0.34*	-	*-0.51*	0.31	0.28	0.16	*0.36*	0.25	-0.18	0.13	0.16
ExtLowDur	0.19	0.04	-0.18	-0.18	-0.17	-0.17	-0.28	*0.36*	0.03	-0.32	0.00	0.11	*-0.51*	-	-0.02	*-0.35*	-0.33	-0.23	-0.21	0.12	0.10	-0.32

ExtLow Time	-0.02	-0.11	0.31	0.27	0.26	0.20	*-0.37*	0.02	*0.36*	*-0.38*	-0.21	*0.61*	0.31	-0.02	-	*-0.45*	-0.21	0.22	0.19	0.18	*0.39*	0.00
ExtLow Freq	-0.17	*0.73*	-0.30	-0.24	-0.22	-0.24	*0.96*	-0.29	*-0.41*	*0.64*	-0.26	0.00	0.28	*-0.35*	*-0.45*	-	-0.03	*-0.47*	*-0.37*	*0.41*	0.05	-0.33
High Dur	-0.03	*-0.40*	*0.75*	*0.80*	*0.80*	*0.84*	-0.12	*-0.37*	*0.35*	0.22	*0.58*	*-0.63*	0.16	-0.33	-0.21	-0.03	-	*0.56*	0.20	-0.24	*-0.62*	*0.38*
Small Flood Peak	-0.09	*-0.66*	*0.78*	*0.66*	*0.63*	*0.62*	*-0.42*	-0.24	*0.47*	0.09	0.31	-0.11	*0.36*	-0.23	0.22	*-0.47*	*0.56*	-	*0.85*	*-0.56*	-0.20	*0.78*
Small Flood Rise	-0.18	*-0.51*	*0.43*	0.29	0.25	0.18	-0.32	*-0.37*	0.08	-0.04	0.14	0.05	0.25	-0.21	0.19	*-0.37*	0.20	*0.85*	-	*-0.50*	0.15	*0.86*
Small Flood Fall	*-0.48*	*0.53*	-0.07	0.03	0.05	-0.08	*0.41*	-0.08	-0.03	-0.04	*-0.39*	0.17	-0.18	0.12	0.18	*0.41*	-0.24	*-0.56*	*-0.50*	-	0.23	*-0.63*
Large Flood Dur	-0.34	*0.36*	*-0.35*	*-0.37*	*-0.38*	*-0.46*	0.11	-0.13	*-0.48*	*-0.40*	*-0.54*	*0.54*	0.13	0.10	*0.39*	0.05	*-0.62*	-0.20	0.15	0.23	-	-0.12
Large Flood Rise	0.03	*-0.56*	*0.36*	0.24	0.21	0.19	*-0.42*	*-0.39*	0.02	0.02	0.29	-0.32	0.16	-0.32	0.00	-0.33	*0.38*	*0.78*	*0.86*	*-0.63*	-0.12	-

Table A1.7: RBI and < 1 cfs Combined Calibration flow metric correlation matrix with correlations statistically significant at the p = 0.1 level denoted in italics.

Flow Metric	**RBI_1hr**	**cfs_1hr**	**Dec**	**Jan**	**Feb**	**7DayMax**	**#ZeroDays**	**DateMax**	**LowPulseCount**	**HighPulseCount**	**Rise Rate**	**Fall Rate**	**#Reversals**	**Ext Low Dur**	**Ext Low Time**	**Ext Low Freq**	**High Dur**	**SmallFloodPeak**	**SmallFloodRise**	**SmallFloodFall**	**Large Flood Dur**	**LargeFloodRise**
RBI_1hr	-	-0.13	-0.21	-0.08	*-0.40*	0.05	-0.05	0.13	0.15	*0.55*	-0.27	0.16	0.23	-0.29	0.17	-0.02	-0.05	0.03	-0.16	*-0.44*	*-0.55*	*0.44*
cfs_1hr	-0.13	-	*-0.63*	*-0.66*	*-0.55*	*-0.51*	*0.97*	0.07	*-0.85*	*-0.37*	*-0.40*	*0.37*	*-0.69*	-0.15	*-0.96*	*0.80*	*0.84*	*-0.49*	*-0.43*	0.08	*-0.34*	-0.31
Dec	-0.21	*-0.63*	-	*0.97*	*0.95*	*0.77*	*-0.56*	*-0.35*	*0.43*	-0.17	*0.92*	*-0.90*	*0.37*	0.11	*0.46*	*-0.48*	*-0.36*	*0.76*	*0.78*	0.12	*0.50*	*0.36*
Jan	-0.08	*-0.66*	*0.97*	-	*0.86*	*0.80*	*-0.57*	*-0.41*	*0.49*	-0.14	*0.84*	*-0.90*	*0.49*	0.09	*0.50*	*-0.49*	*-0.40*	*0.78*	*0.64*	0.20	*0.45*	*0.48*
Feb	*-0.40*	*-0.55*	*0.95*	*0.86*	-	*0.72*	*-0.50*	-0.30	*0.37*	-0.22	*0.94*	*-0.85*	0.26	0.17	*0.37*	*-0.43*	-0.28	*0.72*	*0.84*	0.10	*0.51*	0.22
7Day Max	0.05	*-0.51*	*0.77*	*0.80*	*0.72*	-	*-0.47*	-0.18	*0.59*	0.10	*0.75*	*-0.85*	*0.61*	0.10	*0.42*	*-0.38*	-0.24	*0.98*	*0.65*	-0.02	0.10	*0.82*
#Zero Days	-0.05	*0.97*	*-0.56*	*-0.57*	*-0.50*	*-0.47*	-	0.02	*-0.84*	*-0.38*	-0.32	0.30	*-0.73*	-0.22	*-0.97*	*0.79*	*0.83*	*-0.44*	*-0.38*	0.07	*-0.39*	-0.26
Date Max	0.13	0.07	*-0.35*	*-0.41*	-0.30	-0.18	0.02	-	0.01	0.31	*-0.43*	*0.44*	-0.04	-0.03	0.09	0.00	-0.27	-0.21	0.07	*-0.63*	-0.12	-0.04
Low Pulse	0.15	*-0.85*	*0.43*	*0.49*	*0.37*	*0.59*	*-0.84*	0.01	-	*0.46*	0.28	-0.30	*0.77*	0.18	*0.86*	*-0.73*	*-0.72*	*0.58*	0.28	-0.09	0.03	*0.51*

Count																						
High Pulse Count	*0.55*	*-0.37*	-0.17	-0.14	-0.22	0.10	*-0.38*	0.31	*0.46*	-	-0.33	0.25	*0.47*	-0.01	*0.52*	-0.10	-0.33	0.05	0.00	*-0.47*	*-0.41*	0.28
Rise Rate	-0.27	*-0.40*	*0.92*	*0.84*	*0.94*	*0.75*	-0.32	*-0.43*	0.28	-0.33	-	*-0.89*	0.15	0.16	0.18	*-0.36*	-0.07	*0.78*	*0.75*	0.17	0.32	*0.35*
Fall Rate	0.16	*0.37*	*-0.90*	*-0.90*	*-0.85*	*-0.85*	0.30	*0.44*	-0.30	0.25	*-0.89*	-	-0.33	-0.13	-0.21	0.30	0.07	*-0.80*	*-0.69*	-0.23	*-0.39*	*-0.51*
#Reversals	0.23	*-0.69*	*0.37*	*0.49*	0.26	*0.61*	*-0.73*	-0.04	*0.77*	*0.47*	0.15	-0.33	-	0.01	*0.76*	*-0.35*	*-0.53*	*0.58*	0.21	-0.12	0.10	*0.56*
ExtLowDur	-0.29	-0.15	0.11	0.09	0.17	0.10	-0.22	-0.03	0.18	-0.01	0.16	-0.13	0.01	-	0.15	*-0.50*	-0.21	0.09	-0.02	*0.37*	0.09	0.10
ExtLow Time	0.17	*-0.96*	*0.46*	*0.50*	*0.37*	*0.42*	*-0.97*	0.09	*0.86*	*0.52*	0.18	-0.21	*0.76*	0.15	-	*-0.75*	*-0.89*	*0.37*	0.30	-0.14	0.31	0.30
ExtLow Freq	-0.02	*0.80*	*-0.48*	*-0.49*	*-0.43*	*-0.38*	*0.79*	0.00	*-0.73*	-0.10	*-0.36*	0.30	*-0.35*	*-0.50*	*-0.75*	-	*0.80*	*-0.35*	-0.24	-0.16	-0.33	-0.29
High Dur	-0.05	*0.84*	*-0.36*	*-0.40*	-0.28	-0.24	*0.83*	-0.27	*-0.72*	-0.33	-0.07	0.07	*-0.53*	-0.21	*-0.89*	*0.80*	-	-0.20	-0.17	0.05	*-0.36*	-0.16
Small Flood Peak	0.03	*-0.49*	*0.76*	*0.78*	*0.72*	*0.98*	*-0.44*	-0.21	*0.58*	0.05	*0.78*	*-0.80*	*0.58*	0.09	*0.37*	*-0.35*	-0.20	-	*0.63*	-0.04	0.00	*0.80*
Small Flood Rise	-0.16	*-0.43*	*0.78*	*0.64*	*0.84*	*0.65*	-0.38	0.07	0.28	0.00	*0.75*	*-0.69*	0.21	-0.02	0.30	-0.24	-0.17	*0.63*	-	*-0.41*	0.31	0.23
Small Flood Fall	*-0.44*	0.08	0.12	0.20	0.10	-0.02	0.07	*-0.63*	-0.09	*-0.47*	0.17	-0.23	-0.12	*0.37*	-0.14	-0.16	0.05	-0.04	*-0.41*	-	*0.38*	-0.07
Large Flood Dur	*-0.55*	*-0.34*	*0.50*	*0.45*	*0.51*	0.10	*-0.39*	-0.12	0.03	*-0.41*	0.32	*-0.39*	0.10	0.09	0.31	-0.33	*-0.36*	0.00	0.31	*0.38*	-	-0.30
Large Flood Rise	*0.44*	-0.31	*0.36*	*0.48*	0.22	*0.82*	-0.26	-0.04	*0.51*	0.28	*0.35*	*-0.51*	*0.56*	0.10	0.30	-0.29	-0.16	*0.80*	0.23	-0.07	-0.30	-

Appendix II: Correlation matrices for all eight biotic metrics with their sets of reduced flow metrics

Table A2.1: EPTPercentTaxa correlation matrix using its reduced set of flow metrics for all gage and calibrated model data. Correlations statistically significant at the p = 0.1 level are denoted in italics and differing correlation directions relative to the 3 WY antecedent gage data are underlined.

Flow Metric	3 WY Antecedent Gage	WY 2005-2007 Gage	NSE Calibration	RBI Calibration	< 1 cfs Calibration	RBI and < 1 cfs Combined Calibration
RBI_1hr	0.23	0.08	0.16	0.08	<u>-0.13</u>	0.12
cfs_1hr	*-0.35*	-0.21	<u>*0.50*</u>	-0.09	-0.20	-0.21
NumZeroDays	*-0.44*	-0.34	<u>*0.55*</u>	-0.01	*-0.37*	-0.30
NumReversals	*0.40*	0.24	<u>-0.26</u>	<u>*-0.35*</u>	<u>-0.07</u>	*0.58*
ExtremeLowDuration	*-0.43*	*-0.38*	<u>0.09</u>	<u>0.15</u>	-0.12	<u>0.19</u>
ExtremeLowTiming	*0.57*	*0.40*	<u>*-0.55*</u>	<u>-0.02</u>	0.14	*0.37*
HighFlowDuration	*-0.44*	<u>*0.48*</u>	<u>0.16</u>	<u>0.23</u>	<u>0.24</u>	*-0.35*
LargeFloodDuration	*-0.41*	-0.30	-0.01	-0.26	-0.08	<u>0.13</u>

Table A2.2: DesiResist correlation matrix using its reduced set of flow metrics for all gage and calibrated model data. Correlations statistically significant at the p = 0.1 level are denoted in italics and differing correlation directions relative to the 3 WY antecedent gage data are underlined.

Flow Metric	3 WY Antecedent Gage	WY 2005-2007 Gage	NSE Calibration	RBI Calibration	< 1 cfs Calibration	RBI and < 1 cfs Combined Calibration
RBI_1hr	-0.16	-0.13	-0.30	-0.13	<u>0.04</u>	-0.17
cfs_1hr	*0.54*	*0.48*	<u>*-0.37*</u>	*0.36*	*0.47*	*0.48*
NumZeroDays	*0.57*	*0.55*	<u>*-0.40*</u>	0.34	*0.56*	*0.53*
LowPulseCount	*-0.42*	-0.32	-0.04	-	-0.26	*-0.49*
NumReversals	*-0.54*	*-0.45*	<u>0.19</u>	<u>0.26</u>	<u>0.20</u>	*-0.69*
ExtremeLowDuration	*0.48*	*0.57*	0.08	-0.06	0.09	<u>-0.27</u>
ExtremeLowTiming	*-0.68*	*-0.58*	<u>*0.38*</u>	-0.06	-0.16	*-0.58*
ExtremeLowFreq	*-0.35*	*-0.53*	<u>*0.36*</u>	<u>0.19</u>	<u>*0.43*</u>	<u>0.30</u>
HighFlowDuration	*0.56*	<u>-0.31</u>	<u>-0.24</u>	<u>-0.22</u>	<u>*-0.38*</u>	*0.47*
SmallFloodRiserate	*-0.35*	-0.26	-0.19	0.05	0.13	<u>0.13</u>
LargeFloodDuration	*0.47*	*0.55*	0.01	0.08	0.26	<u>-0.09</u>

Table A2.3: NoninsectTaxa correlation matrix using its reduced set of flow metrics for all gage and calibrated model data. Correlations statistically significant at the p = 0.1 level are denoted in italics and differing correlation directions relative to the 3 WY antecedent gage data are underlined.

Flow Metric	3 WY Antecedent Gage	WY 2005-2007 Gage	NSE Calibration	RBI Calibration	< 1 cfs Calibration	RBI and < 1 cfs Combined Calibration
RBI_1hr	*-0.35*	-0.08	<u>0.10</u>	-0.08	-0.24	-0.04
cfs_1hr	*-0.44*	*-0.50*	-0.07	*-0.55*	*-0.50*	*-0.50*
NumZeroDays	*-0.41*	*-0.45*	-0.19	*-0.60*	*-0.42*	*-0.44*
DateMax	*0.43*	*0.47*	<u>-0.02</u>	0.04	*-0.33*	*-0.37*
ExtremeLowTiming	*0.44*	*0.42*	<u>-0.04</u>	0.15	<u>-0.08</u>	*0.45*
LargeFloodDuration	*-0.35*	*-0.47*	0.02	0.12	-0.20	0.16

Table A2.4: DisturbResil correlation matrix using its reduced set of flow metrics for all gage and calibrated model data. Correlations statistically significant at the p = 0.1 level are denoted in italics and differing correlation directions relative to the 3 WY antecedent gage data are underlined.

Flow Metric	3 WY Antecedent Gage	WY 2005-2007 Gage	NSE Calibration	RBI Calibration	< 1 cfs Calibration	RBI and < 1 cfs Combined Calibration
RBI_1hr	0.07	0.14	0.08	0.14	0.22	0.12
cfs_1hr	0.26	0.25	<u>*-0.46*</u>	0.08	0.24	0.25
ExtremeLowTiming	*-0.57*	-0.31	<u>*0.40*</u>	-0.01	-0.13	*-0.37*
HighFlowDuration	0.32	<u>*-0.46*</u>	<u>-0.21</u>	<u>-0.21</u>	0.00	*0.35*
LargeFloodDuration	*0.39*	0.20	<u>-0.16</u>	0.06	-0.06	<u>-0.13</u>

Table A2.5: AmphipodaPercent correlation matrix using its reduced set of flow metrics for all gage and calibrated model data. Correlations statistically significant at the p = 0.1 level are denoted in italics and differing correlation directions relative to the 3 WY antecedent gage data are underlined.

Flow Metric	3 WY Antecedent Gage	WY 2005-2007 Gage	NSE Calibration	RBI Calibration	< 1 cfs Calibration	RBI and < 1 cfs Combined Calibration
RBI_1hr	*-0.37*	*-0.41*	-0.20	*-0.41*	-0.14	*-0.43*
cfs_1hr	-0.10	-0.32	-0.30	-0.18	-0.33	-0.32
Feb	*0.42*	0.27	0.09	0.15	0.05	*0.72*
RiseRate	*0.51*	0.24	<u>-0.06</u>	<u>-0.12</u>	0.01	*0.64*
FallRate	*-0.36*	-0.18	-0.06	<u>0.04</u>	<u>0.16</u>	*-0.45*

ExtremeLowDuration	*0.57*	-0.21	-0.28	-0.13	-0.11	0.04
LargeFloodRiserate	*0.35*	*0.69*	0.31	*0.70*	*0.64*	-0.18

Table A2.6: SCIBI correlation matrix using its reduced set of flow metrics for all gage and calibrated model data. Correlations statistically significant at the p = 0.1 level are denoted in italics and differing correlation directions relative to the 3 WY antecedent gage data are underlined.

Flow Metric	3 WY Antecedent Gage	WY 2005-2007 Gage	NSE Calibration	RBI Calibration	< 1 cfs Calibration	RBI and < 1 cfs Combined Calibration
RBI_1hr	-0.14	-0.21	-0.11	-0.21	-0.02	-0.21
cfs_1hr	0.11	0.13	*0.47*	0.25	0.16	0.13
Dec	-0.33	*-0.34*	-0.20	-0.14	-0.30	*-0.38*
Jan	-0.31	*-0.38*	-0.10	-0.09	-0.23	-0.31
SmallFloodFallrate	*0.34*	0.15	0.11	*0.37*	0.18	0.26

Table A2.7: SndInstabResist correlation matrix using its reduced set of flow metrics for all gage and calibrated model data. Correlations statistically significant at the p = 0.1 level are denoted in italics and differing correlation directions relative to the 3 WY antecedent gage data are underlined.

Flow Metric	3 WY Antecedent Gage	WY 2005-2007 Gage	NSE Calibration	RBI Calibration	< 1 cfs Calibration	RBI and < 1 cfs Combined Calibration
RBI_1hr	0.03	-0.03	-0.08	-0.03	0.22	-0.06
cfs_1hr	0.31	0.22	*-0.45*	0.03	0.20	0.22
NumZeroDays	0.33	0.25	*-0.46*	0.02	0.27	0.27
ExtremeLowDuration	*0.36*	0.27	0.14	-0.22	0.19	-0.20
ExtremeLowTiming	*-0.43*	*-0.35*	*0.56*	0.18	-0.19	-0.33

Table A2.8: ShredderPercentTaxa correlation matrix using its reduced set of flow metrics for all gage and calibrated model data. Correlations statistically significant at the p = 0.1 level are denoted in italics and differing correlation directions relative to the 3 WY antecedent gage data are underlined.

Flow Metric	3 WY Antecedent Gage	WY 2005-2007 Gage	NSE Calibration	RBI Calibration	< 1 cfs Calibration	RBI and < 1 cfs Combined Calibration
RBI_1hr	*-0.52*	*-0.41*	-0.24	*-0.41*	-0.30	*-0.42*
cfs_1hr	0.06	0.02	0.23	0.25	0.04	0.02

SevenDayMax	-0.29	-0.07	-0.18	-0.19	-0.06	-0.25
HighPulseCount	-0.29	-0.28	-0.12	-0.23	*-0.36*	-0.08
SmallFloodPeak	-0.29	-0.16	-0.24	-0.23	-0.01	-0.27
SmallFloodRiserate	*-0.41*	-0.08	-0.16	-0.24	0.10	-0.11

Appendix III: Multiple regression results for all eight biotic metrics using each set of reduced flow metrics and those statistically selected from each set

Table A3.1: EPTPercentTaxa multiple regression results with entire set of reduced flow metrics. Adjusted R^2 values for flow data italicized when more EPTPercentTaxa variance is explained than the 3 WY antecedent gage data and adjusted R^2 values for flow data underlined when less EPTPercentTaxa variance is explained.

Flow Data	Adj. R^2 All Predictors	Adj. R^2 Backward Selection Predictors	Backward Selected Predictors	Adj. R^2 Cp	Cp Selected Predictors	Adj. R^2 AIC	AIC Selected Predictors
3 WY antecedent gage	0.13	0.30	ExtremeLowTiming	0.30	ExtremeLowTiming	0.32	ExtremeLowDuration, ExtremeLowTiming
WY 2005-2007 gage	*0.78*	*0.76*	cfs_1hr, ExtremeLowDuration, LargeFloodDuration, HighFlowDuration	*0.76*	cfs_1hr, ExtremeLowDuration, LargeFloodDuration, HighFlowDuration	*0.79*	RBI_1hr, cfs_1hr, NumZeroDays, NumReversals, ExtremeLowDuration, LargeFloodDuration, HighFlowDuration
NSE	*0.47*	*0.41*	NumZeroDays, ExtremeLowTiming	*0.46*	RBI_1hr, NumZeroDays, ExtremeLowTiming	*0.52*	RBI_1hr, NumZeroDays, NumReversals, ExtremeLowDuration, ExtremeLowTiming, HighFlowDuration
RBI	0.05	0.00	-	0.13	NumReversals, HighFlowDuration	0.13	NumReversals, HighFlowDuration
< 1 cfs	-0.02	0.00	-	0.10	NumZeroDays	0.10	NumZeroDays
RBI and < 1 cfs	*0.48*	*0.49*	cfs_1hr, NumReversals, HighFlowDuration	*0.49*	cfs_1hr, NumReversals, HighFlowDuration	*0.49*	cfs_1hr, NumReversals, HighFlowDuration

Table A3.2: DesiResist multiple regression results with entire set of reduced flow metrics. Adjusted R^2 values for flow data is italicized when more DesiResist variance is explained than the 3 WY antecedent gage data and adjusted R^2 values for flow data is underlined when less DesiResist variance is explained.

Flow Data	Adj. R^2 All Predictors	Adj. R^2 Backward Selection Predictors	Backward Selected Predictors	Adj. R^2 Cp	Cp Selected Predictors	Adj. R^2 AIC	AIC Selected Predictors
3 WY antecedent gage	0.32	0.44	ExtremeLowTiming	0.44	ExtremeLowTiming	0.48	ExtremeLowDuration, ExtremeLowTiming
WY 2005-2007 gage	*0.51*	*0.60*	cfs_1hr, ExtremeLowDuration, ExtremeLowFreq, HighFlowDuration	*0.60*	cfs_1hr, ExtremeLowDuration, ExtremeLowFreq, HighFlowDuration	*0.60*	cfs_1hr, ExtremeLowDuration, ExtremeLowFreq, HighFlowDuration
NSE	*0.52*	*0.63*	LowPulseCount, ExtremeLowFreq, HighFlowDuration	*0.64*	NumReversals, ExtremeLowFreq, HIghFlowDuration	*0.66*	RBI_1hr, LowPulseCount, ExtremeLowFreq, HighFlowDuration, LargeFloodDuration
RBI	*0.44*	*0.46*	RBI_1hr, cfs_1hr, NumZeroDays, NumReversals, ExtremeLowTiming, ExtremeLowFreq, SmallFloodRiserate, LargeFloodDuration	*0.46*	RBI_1hr, cfs_1hr, NumZeroDays, NumReversals, ExtremeLowTiming, ExtremeLowFreq, SmallFloodRiserate, LargeFloodDuration	0.47	RBI_1hr, cfs_1hr, NumZeroDays, NumReversals, ExtremeLowDuration, ExtremeLowTiming, ExtremeLowFreq, SmallFloodRiserate, LargeFloodDuration
< 1 cfs	*0.46*	*0.50*	NumZeroDays, ExtremeLowDuration, SmallFloodRiserate	*0.59*	RBI_1hr, NumZeroDays, ExtremeLowDuration, HighFlowDuration, SmallFloodRiserate	*0.59*	RBI_1hr, NumZeroDays, ExtremeLowDuration, HighFlowDuration, SmallFloodRiserate
RBI and < 1 cfs	*0.50*	*0.46*	NumReversals	*0.57*	NumReversals, ExtremeLowDuration, SmallFloodRiserate	*0.57*	NumReversals, ExtremeLowDuration, SmallFloodRiserate

Table A3.3: NoninsectTaxa multiple regression results with entire set of reduced flow metrics. Adjusted R^2 values for flow data is italicized when more NoninsectTaxa variance is explained than the 3 WY antecedent gage data and adjusted R^2 values for flow data is underlined when less NoninsectTaxa variance is explained.

Flow Data	Adj. R^2 All Predictors	Adj. R^2 Backward Selection Predictors	Backward Selected Predictors	Adj. R^2 Cp	Cp Selected Predictors	Adj. R^2 AIC	AIC Selected Predictors
3 WY antecedent gage	0.30	0.36	RBI_1hr, DateMax, LargeFloodDuration	0.36	RBI_1hr, DateMax, LargeFloodDuration	0.36	RBI_1hr, DateMax, LargeFloodDuration
WY 2005-2007 gage	0.25	0.22	cfs_1hr	0.31	NumZeroDays, DateMax	0.31	NumZeroDays, DateMax

NSE	-0.19	0.00	-	0.00	NumZeroDays	0.00	NumZeroDays
RBI	*0.46*	0.34	NumZeroDays	*0.47*	RBI_1hr, cfs_1hr, NumZeroDays, DateMax, ExtremeLowTiming	*0.47*	RBI_1hr, cfs_1hr, NumZeroDays, DateMax, ExtremeLowTiming
< 1 cfs	*0.38*	0.31	NumZeroDays, DateMax	*0.40*	RBI_1hr, NumZeroDays, DateMax, ExtremeLowTiming	*0.40*	RBI_1hr, NumZeroDays, DateMax, ExtremeLowTiming
RBI and < 1 cfs	*0.31*	0.31	DateMax, ExtremeLowTiming	0.31	DateMax, ExtremeLowTiming	0.31	DateMax, ExtremeLowTiming

Table A3.4: DisturbResil multiple regression results with entire set of reduced flow metrics. Adjusted R^2 values for flow data is italicized when more DisturbResil variance is explained than the 3 WY antecedent gage data and adjusted R^2 values for flow data is underlined when less DisturbResil variance is explained.

Flow Data	Adj. R^2 All Predictors	Adj. R^2 Backward Selection Predictors	Backward Selected Predictors	Adj. R^2 Cp	Cp Selected Predictors	Adj. R^2 AIC	AIC Selected Predictors
3 WY antecedent gage	0.22	0.29	ExtremeLowTiming	0.29	ExtremeLowTiming	0.29	ExtremeLowTiming
WY 2005-2007 gage	*0.32*	*0.36*	ExtremeLowTiming, HighFlowDuration	*0.36*	ExtremeLowTiming, HighFlowDuration	*0.36*	ExtremeLowTiming, HighFlowDuration
NSE	0.20	*0.29*	cfs_1hr, HighFlowDuration	*0.29*	cfs_1hr, HighFlowDuration	*0.29*	cfs_1hr, HighFlowDuration
RBI	-0.07	0.00	-	0.00	HighFlowDuration	0.04	RBI_1hr, HighFlowDuration
< 1 cfs	-0.08	0.00	-	0.02	cfs_1hr	0.02	cfs_1hr
RBI and < 1 cfs	0.17	0.00	-	0.21	cfs_1hr, ExtremeLowTiming	0.21	cfs_1hr, ExtremeLowTiming

Table A3.5: log(AmphipodaPercent) multiple regression results with entire set of reduced flow metrics. Adjusted R^2 values for flow data is italicized when more log(AmphipodaPercent) variance is explained than the 3 WY antecedent gage data and adjusted R^2 values for flow data is underlined when less log(AmphipodaPercent) variance is explained.

Flow Data	Adj. R^2 All Predictors	Adj. R^2 Backward Selection Predictors	Backward Selected Predictors	Adj. R^2 Cp	Cp Selected Predictors	Adj. R^2 AIC	AIC Selected Predictors
3 WY antecedent gage	0.19	0.28	log(LargeFloodRiseRate)	0.28	log(LargeFloodRiseRate)	0.28	log(LargeFloodRiseRate)
WY 2005-2007 gage	*0.34*	0.27	cfs_1hr	*0.38*	cfs_1hr, Feb, FallRate	*0.38*	cfs_1hr, Feb, FallRate
NSE	0.04	0.00	-	0.17	RBI_1hr, cfs_1hr	0.17	RBI_1hr, cfs_1hr
RBI	*0.46*	*0.50*	log(RiseRate), ExtremeLowDuration, FallRate,	*0.50*	log(RiseRate), ExtremeLowDuration, FallRate,	*0.50*	log(RiseRate), ExtremeLowDuration, FallRate,

			log(LargeFloodRiseRate)		log(LargeFloodRiseRate)		log(LargeFloodRiseRate)
< 1 cfs	*0.34*	0.27	cfs_1hr	*0.40*	cfs_1hr, FallRate, log(LargeFloodRiseRate)	*0.43*	cfs_1hr, log(RiseRate), FallRate, log(LargeFloodRiseRate)
RBI and < 1 cfs	*0.53*	*0.59*	RBI_1hr, Feb, FallRate	*0.59*	RBI_1hr, Feb, FallRate	*0.59*	RBI_1hr, Feb, FallRate

Table A3.6: SCIBI multiple regression results with entire set of reduced flow metrics. Adjusted R^2 values for flow data is italicized when more SCIBI variance is explained than the 3 WY antecedent gage data and adjusted R^2 values for flow data is underlined when less SCIBI variance is explained.

Flow Data	Adj. R^2 All Predictors	Adj. R^2 Backward Selection Predictors	Backward Selected Predictors	Adj. R^2 Cp	Cp Selected Predictors	Adj. R^2 AIC	AIC Selected Predictors
3 WY antecedent gage	-0.09	0.00	-	0.08	SmallFloodFallrate	0.08	SmallFloodFallrate
WY 2005-2007 gage	*0.16*	0.00	-	*0.10*	Jan	*0.10*	Jan
NSE	*0.23*	*0.28*	cfs_1hr, Dec, Jan	*0.28*	cfs_1hr, Dec, Jan	*0.28*	cfs_1hr, Dec, Jan
RBI	*0.10*	0.00	-	*0.10*	SmallFloodFallrate	*0.10*	SmallFloodFallrate
< 1 cfs	*0.00*	0.00	-	*0.12*	Dec, Jan	*0.12*	Dec, Jan
RBI and < 1 cfs	*0.33*	*0.32*	RBI_1hr, Dec, Jan	*0.32*	RBI_1hr, Dec, Jan	*0.35*	RBI_1hr, cfs_1hr, Dec, Jan

TableA3.7: SndInstabResist multiple regression results with entire set of reduced flow metrics. Adjusted R^2 values for flow data is italicized when more SndInstabResist variance is explained than the 3 WY antecedent gage data and adjusted R^2 values for flow data is underlined when less SndInstabResist variance is explained.

Flow Data	Adj. R^2 All Predictors	Adj. R^2 Backward Selection Predictors	Backward Selected Predictors	Adj. R^2 Cp	Cp Selected Predictors	Adj. R^2 AIC	AIC Selected Predictors
3 WY antecedent gage	0.15	0.15	ExtremeLowTiming	0.15	ExtremeLowTiming	0.15	ExtremeLowTiming
WY 2005-2007 gage	-0.10	0.00	-	0.08	ExtremeLowTiming	0.08	ExtremeLowTiming
NSE	*0.33*	*0.29*	ExtremeLowTiming	*0.35*	NumZeroDays, ExtremeLowTiming	*0.35*	NumZeroDays, ExtremeLowTiming
RBI	-0.05	0.00	-	0.01	ExtremeLowDuration	0.01	ExtremeLowDuration
< 1 cfs	0.02	0.00	-	0.03	NumZeroDays	0.07	RBI_1hr, NumZeroDays
RBI and < 1 cfs	0.07	0.00	-	*0.17*	cfs_1hr, ExtremeLowTiming	*0.17*	cfs_1hr, ExtremeLowTiming

Table A3.8: ShredderPercentTaxa multiple regression results with entire set of reduced flow metrics. Adjusted R^2 values for flow data is italicized when more ShredderPercentTaxa variance is explained than the 3 WY antecedent gage data and adjusted R^2 values for flow data is underlined when less ShredderPercentTaxa variance is explained.

Flow Data	Adj. R^2 All Predictors	Adj. R^2 Backward Selection Predictors	Backward Selected Predictors	Adj. R^2 Cp	Cp Selected Predictors	Adj. R^2 AIC	AIC Selected Predictors
3 WY antecedent gage	0.14	0.24	RBI_1hr	0.24	RBI_1hr	0.24	RBI_1hr
WY 2005-2007 gage	-0.01	0.00	-	0.13	RBI_1hr	0.13	RBI_1hr
NSE	-0.02	0.00	-	0.08	SevenDayMax, SmallFloodPeak	0.08	SevenDayMax, SmallFloodPeak
RBI	0.02	0.13	RBI_1hr	0.13	RBI_1hr	0.13	RBI_1hr
< 1 cfs	-0.05	0.00	-	0.09	HighPulseCount	0.09	HighPulseCount
RBI and < 1 cfs	0.09	0.14	RBI_1hr	0.14	RBI_1hr	0.18	RBI_1hr, SmallFloodPeak

Appendix IV: Multiple regression results for all eight biotic metrics using only three flow metrics and those statistically selected from the three

Table A4.1: EPTPercentTaxa multiple regression results with further reduced set of 3 flow metrics. Adjusted R^2 values for flow data is italicized when more EPTPercentTaxa variance is explained than the 3 WY antecedent gage data and adjusted R^2 values for flow data is underlined when less EPTPercentTaxa variance is explained.

Flow Data	Adj. R^2 All Predictors	Adj. R^2 Backward Selection Predictors	Backward Selected Predictors	Adj. R^2 Cp	Cp Selected Predictors	Adj. R^2 AIC	AIC Selected Predictors
3 WY antecedent gage	0.32	0.30	ExtremeLowTiming	0.30	ExtremeLowTiming	0.32	ExtremeLowDuration, ExtremeLowTiming
WY 2005-2007 gage	0.04	0.00	-	0.12	ExtremeLowTiming	0.12	ExtremeLowTiming
NSE	*0.39*	*0.41*	NumZeroDays, ExtremeLowTiming	*0.41*	NumZeroDays, ExtremeLowTiming	*0.41*	NumZeroDays, ExtremeLowTiming
RBI	-0.11	0.00	-	-0.02	ExtremeLowDuration	-0.02	ExtremeLowDuration
< 1 cfs	0.08	0.00	-	0.10	NumZeroDays	0.10	NumZeroDays
RBI and < 1 cfs	0.13	0.00	-	0.10	ExtremeLowTiming	0.10	ExtremeLowTiming

Table A4.2: DesiResist multiple regression results with further reduced set of 3 flow metrics. Adjusted R^2 values for flow data is italicized when more DesiResist variance is explained than the 3 WY antecedent gage data and adjusted R^2 values for flow data is underlined when less DesiResist variance is explained.

Flow Data	Adj. R^2 All Predictors	Adj. R^2 Backward Selection Predictors	Backward Selected Predictors	Adj. R^2 Cp	Cp Selected Predictors	Adj. R^2 AIC	AIC Selected Predictors
3 WY antecedent gage	0.40	0.44	ExtremeLowTiming	0.44	ExtremeLowTiming	0.44	ExtremeLowTiming
WY 2005-2007 gage	0.27	0.31	ExtremeLowTiming	0.31	ExtremeLowTiming	0.31	ExtremeLowTiming
NSE	0.13	0.12	NumZeroDays	0.16	NumZeroDays, ExtremeLowTiming	0.16	NumZeroDays, ExtremeLowTiming
RBI	0.01	0.00	-	0.09	cfs_1hr	0.09	cfs_1hr
< 1 cfs	0.22	0.28	NumZeroDays	0.28	NumZeroDays	0.28	NumZeroDays
RBI and < 1 cfs	0.35	0.31	ExtremeLowTiming	0.38	cfs_1hr, ExtremeLowTiming	0.38	cfs_1hr, ExtremeLowTiming

Table A4.3: NoninsectTaxa multiple regression results with further reduced set of 3 flow metrics. Adjusted R^2 values for flow data is italicized when more NoninsectTaxa variance is explained than the 3 WY antecedent gage data and adjusted R^2 values for flow data is underlined when less NoninsectTaxa variance is explained.

Flow Data	Adj. R^2 All Predictors	Adj. R^2 Backward Selection Predictors	Backward Selected Predictors	Adj. R^2 Cp	Cp Selected Predictors	Adj. R^2 AIC	AIC Selected Predictors
3 WY antecedent gage	0.12	0.16	cfs_1hr	0.16	cfs_1hr	0.16	cfs_1hr
WY 2005-2007 gage	*0.17*	*0.22*	cfs_1hr	*0.22*	cfs_1hr	*0.22*	cfs_1hr
NSE	-0.08	0.00	-	0.00	NumZeroDays	0.00	NumZeroDays
RBI	*0.37*	*0.34*	NumZeroDays	*0.34*	NumZeroDays	*0.34*	NumZeroDays
< 1 cfs	*0.18*	*0.21*	cfs_1hr	*0.21*	cfs_1hr	*0.21*	cfs_1hr
RBI and < 1 cfs	*0.19*	*0.22*	cfs_1hr	*0.22*	cfs_1hr	*0.22*	cfs_1hr

Table A4.4: DisturbResist multiple regression results with further reduced set of 3 flow metrics. Adjusted R^2 values for flow data is italicized when more DisturbResist variance is explained than the 3 WY antecedent gage data and adjusted R^2 values for flow data is underlined when less DisturbResist variance is explained.

Flow Data	Adj. R^2 All Predictors	Adj. R^2 Backward Selection Predictors	Backward Selected Predictors	Adj. R^2 Cp	Cp Selected Predictors	Adj. R^2 AIC	AIC Selected Predictors
3 WY antecedent gage	0.23	0.29	ExtremeLowTiming	0.29	ExtremeLowTiming	0.29	ExtremeLowTiming

WY 2005-2007 gage	*0.35*	*0.36*	ExtremeLowTiming, HighFlowDuration	*0.36*	ExtremeLowTiming, HighFlowDuration	*0.36*	ExtremeLowTiming, HighFlowDuration
NSE	0.15	0.13	ExtremeLowTiming	0.17	ExtremeLowTiming, LargeFloodDuration	0.17	ExtremeLowTiming, LargeFloodDuration
RBI	-0.08	0.00	-	0.00	HighFlowDuration	0.00	HighFlowDuration
< 1 cfs	-0.12	0.00	-	-0.03	ExtremeLowTiming	-0.03	ExtremeLowTiming
RBI and < 1 cfs	0.02	0.00	-	0.10	ExtremeLowTiming	0.10	ExtremeLowTiming

Table A4.5: log(Amphipoda) multiple regression results with further reduced set of 3 flow metrics. Adjusted R^2 values for flow data is italicized when more log(Amphipoda) variance is explained than the 3 WY antecedent gage data and adjusted R^2 values for flow data is underlined when less log(Amphipoda) variance is explained.

Flow Data	Adj. R^2 All Predictors	Adj. R^2 Backward Selection Predictors	Backward Selected Predictors	Adj. R^2 Cp	Cp Selected Predictors	Adj. R^2 AIC	AIC Selected Predictors
3 WY antecedent gage	0.10	0.16	log(RiseRate)	0.16	log(RiseRate)	0.16	log(RiseRate)
WY 2005-2007 gage	*0.24*	*0.27*	log(RiseRate)	*0.27*	log(RiseRate)	*0.27*	log(RiseRate)
NSE	0.08	0.00	-	0.04	ExtremeLowDuration	0.04	ExtremeLowDuration
RBI	*0.32*	*0.24*	log(RiseRate)	*0.32*	log(RiseRate), ExtremeLowDuration	*0.32*	log(RiseRate), ExtremeLowDuration
< 1 cfs	*0.14*	*0.19*	log(RiseRate)	*0.19*	log(RiseRate)	*0.19*	log(RiseRate)
RBI and < 1 cfs	*0.37*	*0.38*	Feb	*0.38*	Feb	*0.38*	Feb

Table A4.6: SCIBI multiple regression results with further reduced set of 3 flow metrics. Adjusted R^2 values for flow data is italicized when more SCIBI variance is explained than the 3 WY antecedent gage data and adjusted R^2 values for flow data is underlined when less SCIBI variance is explained.

Flow Data	Adj. R^2 All Predictors	Adj. R^2 Backward Selection Predictors	Backward Selected Predictors	Adj. R^2 Cp	Cp Selected Predictors	Adj. R^2 AIC	AIC Selected Predictors
3 WY antecedent gage	0.01	0.00	-	0.08	SmallFloodFallrate	0.08	SmallFloodFallrate
WY 2005-2007 gage	*0.06*	0.00	-	*0.10*	Jan	*0.10*	Jan
NSE	*0.10*	0.00	-	*0.10*	Dec, Jan, SmallFloodFallrate	*0.10*	Dec, Jan, SmallFloodFallrate
RBI	*0.03*	0.00	-	*0.10*	SmallFloodFallrate	*0.10*	SmallFloodFallrate
< 1 cfs	*0.07*	0.00	-	*0.12*	Dec, Jan	*0.12*	Dec, Jan

RBI and < 1 cfs	*0.14*	0.00	-	*0.17*	Dec, SmallFloodFallrate	*0.17*	Dec, SmallFloodFallrate

Table A4.7: SndInstabResist multiple regression results with further reduced set of 3 flow metrics. Adjusted R^2 values for flow data is italicized when more SndInstabResist variance is explained than the 3 WY antecedent gage data and adjusted R^2 values for flow data is underlined when less SndInstabResist variance is explained.

Flow Data	Adj. R^2 All Predictors	Adj. R^2 Backward Selection Predictors	Backward Selected Predictors	Adj. R^2 Cp	Cp Selected Predictors	Adj. R^2 AIC	AIC Selected Predictors
3 WY antecedent gage	0.14	0.15	ExtremeLowTiming	0.15	ExtremeLowTiming	0.15	ExtremeLowTiming
WY 2005-2007 gage	<u>0.00</u>	<u>0.00</u>	-	<u>0.08</u>	ExtremeLowTiming	<u>0.08</u>	ExtremeLowTiming
NSE	*0.35*	*0.29*	ExtremeLowTiming	*0.35*	NumZeroDays, ExtremeLowTiming	*0.35*	NumZeroDays, ExtremeLowTiming
RBI	<u>-0.02</u>	<u>0.00</u>	-	<u>0.01</u>	ExtremeLowDuration	<u>0.01</u>	ExtremeLowDuration
< 1 cfs	<u>0.03</u>	<u>0.00</u>	-	<u>0.03</u>	NumZeroDays	<u>0.07</u>	NumZeroDays, ExtremeLowDuration
RBI and < 1 cfs	<u>0.11</u>	<u>0.00</u>	-	<u>0.07</u>	ExtremeLowTiming	<u>0.07</u>	ExtremeLowTiming

Table A4.8: ShredderPercentTaxa multiple regression results with further reduced set of 3 flow metrics. Adjusted R^2 values for flow data is italicized when more ShredderPercentTaxa variance is explained than the 3 WY antecedent gage data and adjusted R^2 values for flow data is underlined when less ShredderPercentTaxa variance is explained.

Flow Data	Adj. R^2 All Predictors	Adj. R^2 Backward Selection Predictors	Backward Selected Predictors	Adj. R^2 Cp	Cp Selected Predictors	Adj. R^2 AIC	AIC Selected Predictors
3 WY antecedent gage	0.25	0.24	RBI_1hr	0.24	RBI_1hr	0.24	RBI_1hr
WY 2005-2007 gage	<u>0.05</u>	<u>0.13</u>	RBI_1hr	<u>0.13</u>	RBI_1hr	<u>0.13</u>	RBI_1hr
NSE	<u>-0.05</u>	<u>0.00</u>	-	<u>0.02</u>	RBI_1hr	<u>0.02</u>	RBI_1hr
RBI	<u>0.11</u>	<u>0.13</u>	RBI_1hr	<u>0.13</u>	RBI_1hr	<u>0.13</u>	RBI_1hr
< 1 cfs	<u>0.08</u>	<u>0.00</u>	-	<u>0.09</u>	HighPulseCount	<u>0.09</u>	HighPulseCount
RBI and < 1 cfs	<u>0.14</u>	<u>0.14</u>	RBI_1hr	<u>0.14</u>	RBI_1hr	<u>0.14</u>	RBI_1hr

Appendix V: Multiple regression results summary

Table A5.1: Multiple regression summary indicating the percent of MRA models in which each set of flow data explained more or less biotic variance than the 3 WY antecedent gage data for the entire set of reduced flow metrics and the further reduced set of 3 flow metrics.

Flow Data	Reduced predictor set		3 predictor set		Reduced and 3 predictor set	
	More	Less	More	Less	More	Less
WY 2005-2007 gage	63%	38%	50%	50%	56%	44%
NSE	50%	50%	38%	63%	44%	56%
RBI	50%	50%	38%	63%	44%	56%
< 1 cfs	50%	50%	38%	63%	44%	56%
RBI and < 1 cfs	63%	38%	38%	63%	50%	50%

Table A5.2: Multiple regression summary indicating the percent of MRA models in which each set of flow data explained more or less biotic variance than the 3 WY antecedent gage data for the AIC selected predictors using the entire set of reduced flow metrics and the further reduced set of 3 flow metrics.

Flow Data	AIC reduced predictor set		AIC 3 predictor set		AIC reduced and 3 predictor set	
	More	Less	More	Less	More	Less
WY 2005-2007 gage	63%	38%	50%	50%	56%	44%
NSE	63%	38%	38%	63%	50%	50%
RBI	38%	63%	38%	63%	38%	63%
< 1 cfs	50%	50%	38%	63%	44%	56%
RBI and < 1 cfs	63%	38%	38%	63%	50%	50%

Appendix VI: Random forest results summary

Table A6.1: Random forest summary indicating the percent of RF models in which each set of flow data explains more or less biotic variance than the 3 WY antecedent gage data for the set of reduced flow metrics.

Flow Data	Reduced predictor set	
	More	Less
WY 2005-2007 gage	88%	13%
NSE	88%	13%
RBI	63%	38%
< 1 cfs	88%	13%
RBI and < 1 cfs	63%	38%

LIST OF ABBREVIATIONS

Acronyms:

AIC	Akaike Information Criterion
CA	California
CDEC	California Data Exchange Center
CIMIS	California Irrigation Management Information System
DEM	digital elevation model
EFC	environmental flow component
ELOHA	Ecological Limits of Hydrologic Alteration
EPA	United States Environmental Protection Agency
EPT	Ephemeroptera, Plecoptera, and Tricoptera
FEMA	Federal Emergency Management Agency
HEC-HMS	Hydrologic Engineering Center Hydrologic Modeling System
IHA	Indicators of Hydrologic Alteration
MRA	multiple linear regression
NATHAT	National Hydrologic Assessment Tool
NCDC	National Climatic Data Center
NED	National Elevation Dataset
NLCD	National Land Cover Database
NOAA	National Oceanic and Atmospheric Administration
NSE	Nash-Sutcliffe Efficiency
PCA	principal component analysis

pers. comm.	personal communication
QA	quality assurance
QC	quality control
RBI	Richards-Baker Flashiness Index
RF	random forest
SC-IBI	Southern California Index of Biotic Integrity
SDCFCD	San Diego County Flood Control District
SMC	Stormwater Monitoring Coalition
SWAMP	Surface Water Ambient Monitoring Program
TNC	The Nature Conservancy
TOC	time of concentration
US/U.S.	United States
USDA	United States Department of Agriculture
USGS	United States Geological Survey
VCWPD	Ventura County Watershed Protection District
WY	Water Year

<u>Metric Abbreviations</u>:

cfs_1hr	percent of time with flow less than 1 cfs using hourly time steps
Adj.	adjusted
Dec.	December
Desi.	desiccation
Dist.	disturbance

Feb.	February
Freq.	frequency
Instab.	instability
Jan.	January
Max.	maximum
Min.	minimum
Num.	number
RBI_1hr	Richards-Baker Flashiness Index using hourly time steps
Resil	resilience
Resist	resistance
SC-IBI	Southern California Index of Biotic Integrity
Snd.	sand
T_{Qmean}	fraction of the record above average flow for the record