

Beyond metrics?

The role of hydrologic baseline archetypes in environmental water management

Belize A Lane^{1*}, Samuel Sandoval-Solis², Eric D Stein³, Sarah M Yarnell⁴, Gregory B Pasternack² and Helen E Dahlke²

*Corresponding author

Email: belize.lane@usu.edu

¹ Department of Civil and Environmental Engineering, Utah State University, Logan, Utah USA

² Department of Land, Air and Water Resources, University of California Davis, California USA

³ Southern California Coastal Water Resources Project, Costa Mesa, California USA

⁴ Center for Watershed Sciences, University of California Davis, California USA

Cite as:

Lane, B.A., Sandoval-Solis, S., Stein, E.D., Yarnell, S.M., Pasternack, G.B. and Dahlke, H.E. 2018. Beyond metrics? The role of hydrologic baseline archetypes in environmental water management. *Environmental Management*. DOI: 10.1007/s00267-018-1077-7

Abstract

Balancing ecological and human water needs often requires characterizing key aspects of the natural flow regime and then predicting ecological response to flow alterations. Flow metrics are generally relied upon to characterize long-term average statistical properties of the natural flow regime (hydrologic baseline conditions). However, some key aspects of hydrologic baseline conditions may be better understood through more complete consideration of continuous patterns of daily, seasonal, and inter-annual variability than through summary metrics. Here, we propose the additional use of high-resolution dimensionless archetypes of regional stream classes to improve understanding of baseline hydrologic conditions and inform regional environmental flows assessments. In an application to California, we describe the development and analysis of hydrologic baseline archetypes to characterize patterns of flow variability within and between stream classes. We then assess the utility of archetypes to provide context for common flow metrics and improve understanding of linkages between aquatic patterns and processes and their hydrologic controls. Results indicate that these archetypes may offer a distinct and complementary tool for researching mechanistic flow-ecology relationships, assessing regional patterns for streamflow management, or understanding impacts of changing climate.

1. Introduction

Managing altered hydrology is a common challenge worldwide in terms of protecting sensitive biological communities (Poff et al. 2010; Bunn and Arthington 2002; Poff and Zimmerman 2010) and balancing the needs of ecological and human water uses (Lane et al. 2014; Postel and Richter 2004). Relationships between flow alteration and ecological characteristics for different natural hydrologic regimes, or *stream classes*, constitute a key element linking the hydrologic, ecological and social aspects of environmental flow assessment (Poff et al. 2010). A critical step in developing these relationships at the regional scale is to identify the degree of hydrologic alteration. This is accomplished by comparing existing (or anticipated future) conditions to a reference condition, or hydrologic baseline, often represented by minimally impacted (e.g., historic) conditions. The *hydrologic baseline* of a stream class is defined here as the (20th century) unimpaired continuous inter- and intra-annual streamflow patterns characteristic for that region and stream class.

Hydrologic classification is a common approach for distinguishing regional hydrologic baselines (Olden et al. 2012). Many such classifications have been developed within the United States (Reidy Liermann et al. 2012; Leibowitz et al. 2016; Wolock et al. 2004) and worldwide (Kennard et al. 2010; Brown et al. 2014b), reflecting an increasing desire to organize and manage complex, highly altered riverscapes to maintain ecosystem integrity. Existing classification systems generally organize natural streamflow patterns by identifying similar streams according to a set of diagnostic flow metrics. When combined with landscape analysis, hydrologic classification can provide a spatially explicit understanding of natural hydrologic variation among streams within and between regions (Wagner 2007). Robust hydrologic classification is expected to improve the development of flow-ecology relationships by reducing the noise in

relational models associated with normal landscape variability. Different hydrologic baselines (i.e., for different stream classes) are expected to respond differently to flow alteration (i.e., they may be more or less sensitive to different forms of alteration), thus providing spatially discontinuous groups of stream reaches that can be managed similarly. In addition to distinguishing dominant hydrologic regimes through classification, establishing regional hydrologic baselines for environmental flows assessments requires subsequent characterization of each hydrologic regime.

Flow metrics are generally relied upon to both distinguish stream classes and then characterize the hydrologic regimes of each class (Poff and Allan 1995; Arthington 2006; Richter 1996). Describing statistical properties related to the magnitude, frequency, duration, timing, and rate of change of streamflow, flow metrics are generally calculated directly from unimpaired flow records (Poff et al. 1997), modelled using watershed attributes (Eng et al. 2017; Parajka et al. 2013), or estimated using flow data from nearby gages (Sawicz et al. 2011). Metrics are often selected for their presumed ecological relevance (Olden et al. 2012) or utility for particular environmental management end-goals (Olden and Poff 2003). For example, in two distinct hydrologic classifications of California, Pyne et al. (2017) sought to inform the ecological limits of alteration while Lane et al. (2017) sought to characterize natural patterns of variability. Towards these distinct end-goals, Pyne et al. (2017) considered average and extreme flow conditions for their established sensitivity to human alteration, while Lane et al. (2017) focused on variability-driven metrics such as median monthly flow and high flow timing to distinguish distinct seasonal and inter-annual patterns thought to be relevant to native biota. These contrasting classifications demonstrate how *a priori* selection of flow metrics for characterizing baseline conditions is influenced by application end-goals (Olden et al. 2012). Further, neither set of metrics provides a continuous characterization of hydrologic baseline conditions capable of capturing the full suite of hydrologic drivers influencing ecosystem response.

Some drivers of ecosystem response may be better understood through more complete consideration of continuous, high-resolution streamflow patterns (Stewart-Koster et al. 2014). Patterns of hydrologic variability, responding to different annual, seasonal, and storm-event scale climate drivers, are layered together to create complex flow regimes that promote variability in channel habitats (Beechie et al. 2010) and contribute to long-term ecological diversity (Naiman et al. 2008). In Mediterranean-montane climates for example, native species are highly adapted to the annually recurring patterns of high and low flows (Lytle and Poff 2004; Gasith and Resh 1999). Differences in the transition from winter flood flows to summer low flows in terms of seasonal predictability (i.e. inter-annual variability) and flashiness between streams have major implications for the life-history strategies of native species. For example, the predictable spring snowmelt recession in mid-elevation snowmelt-driven streams creates favorable habitat conditions and distinct hydrologic cues for spawning and migration of native species (Yarnell et al. 2010). Yet, few standard flow metrics capture the sub-seasonal variability patterns specific to this crucial transition period. Those metrics that do describe critical sub-seasonal patterns are often poorly predicted, limiting their utility at ungaged locations (Eng et al. 2017). We hypothesize that consideration of continuous, daily patterns of hydrologic variability at multiple

temporal scales, in addition to flow metrics, will inform the development of stronger linkages with key ecological cues and processes.

In this paper, we propose the additional use of daily-scale hydrologic baseline archetypes as a complementary approach that allows researchers to more objectively move between conceptual models of flow – ecology relationships, flow metrics, and management applications. Cullum et al. (2017) define an archetype as “a conceptualization of an entire category or class of objects. Archetypes can be framed as abstract exemplars of classes, conceptual models linking form and process and/or tacit mental models similar to those used by field scientists to identify and describe landforms, soils and/or units of vegetation.” Most importantly, an archetype is more than an abstraction of indices, but is a complete realization of an entity. The use of archetypes is well established in hydrology, geomorphology, and biology (McCuen 1989; Parasiewicz et al. 2008; Brown et al. 2014a), but is not common in the ecohydrology literature.

There are several common techniques for generating hydrologic archetypes, reflecting the rich body of literature on synthetic hydrology (Salas 1993), scaling (Blöschl and Sivapalan 1995; Blöschl 2001), and prediction at ungauged basins (Sivapalan 2003), that would be of significant benefit for use in ecohydrology. These techniques involve different levels of complexity, data availability, and applicability, but generally fall into two categories: process-based deterministic watershed models and hydrostatistical approaches (Farmer and Vogel 2013). Watershed modeling is beyond the scope of this study; for a review of regional watershed modeling approaches, see He et al. (2011). Within hydrostatistical approaches, methods range from statistical scaling techniques and fitted probability density function models to more advanced techniques such as autoregressive integrated moving average (ARIMA) models, Fourier analysis (e.g., Pasternack and Hinnov 2003), and wavelets (e.g., Smith et al. 1998; Lundquist and Cayan 2002). Within scaling techniques, streamflow time series can be scaled to represent conditions at different locations or under changing physical controls (e.g., climate change) using scaling variables such as drainage area (Asquith et al. 2006), flow duration curves (Archfield et al. 2007), and average annual flow (Mierau et al. 2017; Farmer and Vogel 2013). Methods that nondimensionalize time series, such as scaling by average annual flow (Yang et al. 2016; Sanborn and Bledsoe 2006; Haines et al. 1988), are particularly useful for establishing baselines, as they improve comparability and hydrologic information transfer across watersheds, regions, storms, etc. [e.g., the dimensionless unit hydrograph for a specific watershed (Bender and Roberson 1961)].

To address the need to establish ecologically relevant hydrologic baselines for the State of California, we propose the development of daily time step, archetypal dimensionless reference hydrographs (DRHs) *in addition* to continued use of flow metrics for quantifying stream class hydrologic baselines. DRHs refer to unimpaired daily streamflow time series nondimensionalized by average annual flow. Once many reference hydrographs are nondimensionalized, subsets of reference gages representing a single stream class can be analyzed together by aggregating their DRHs to create a stream class DRH archetype spanning a

water year. Stream class DRH archetypes can then be used to quantify and compare inter-annual, seasonal, and daily variability patterns within and between stream classes. DRHs characterize the statistical signature and within-class variability of a distinct stream class, but also provide continuous multi-scale information, thus representing a simple, high-resolution hydrologic archetype. In this manner, we expect DRHs to provide context for the selected flow metrics, allowing for connections and relationships between various metrics to be made, thus creating a more complete understanding of flow patterns than the individual metrics alone.

The overall goal of this study is to generate baseline hydrologic archetypes of regional stream classes for California and outline a framework for using these archetypes to inform flow – ecology linkages in support of regional environmental flows assessment. To accomplish this, the specific study objectives are to (1) reconcile two existing hydrologic classifications for California into a single classification using a scientifically defensible reconciliation approach, (2) characterize the resulting stream classes using traditional flow metrics, (3) generate stream class DRH archetypes, (4) characterize the stream classes using the generated DRH archetypes, and (5) evaluate their utility for revealing ecologically significant streamflow patterns. We expect that hydrologic baseline archetypes will provide additional information not available from traditional flow metrics to support process driven analysis of regional streamflow patterns and inform flow - ecology relationships in broad heterogeneous landscapes.

2. Methodology

2.1. Reconciliation of two regional hydrologic classifications

The first phase of this study involved reconciling two existing hydrologic classification systems of the State of California in order to provide a single consensus classification system. These two efforts (Pyne et al. 2017; Lane et al. 2017) addressed different water management objectives, yet arrived at similar, though not totally identical, results. The similarities attest to the resilience of the underlying data in providing a clear signal that drove convergent outcomes. See the *Supplemental Materials* for detailed descriptions of the final classification system.

The reconciliation process consisted of three steps (Figure 1). First, classification geodatabases were merged to identify subregions of agreement and disagreement (Step 1); subregions with disagreements were then resolved through the proposed workflow (Steps 2 and 3). Analysis found that there were 3 major subregions of disagreement, totaling 221,000 km², while the remainder of California (203,000 km²) was generally in agreement. As the Lane et al. (2017) classification provided a more detailed characterization of seasonal and inter-annual hydrologic variability [both aspects of the unimpaired hydrologic regime with well-established ecological significance (Poff et al. 1997)], it was selected as the default classification in subregions of agreement in this study.

Step 2 (Figure 1) assessed the number of reference gages each classification study had in a subregion of disagreement. There were three possible outcomes: the Pyne et al. (2017) classification could have (a) fewer, (b) a similar number, or (c) more reference gages than the

Lane et al. (2017a) classification in the subregions of disagreement. In step 3, different rules were used for reconciliation under each of the three possible outcomes (Figure 1). If there were fewer Pyne et al. (2017) reference gages in a subregion of disagreement, then it was deemed insufficient to warrant creation of a distinct class and the default stream class of Lane et al. (2017) was retained. If there were sufficiently more (>10) reference gages in one subregion where the two classification systems differed in outcome, then it was concluded that there was a reasonable basis to adopt the Pyne et al. (2017) classification, and the additional stream class was incorporated into the reconciled system. Finally, if reference gage quantity and distribution were similar in a subregion of disagreement (Figure 1, outcome b), the hydrologic and physical attributes of the conflicting stream classes were further evaluated to determine if the physical basis and/or statistical patterns were distinct enough to warrant splitting. If the physical distinction was sufficient, the stream classes were split into a sub-class of Lane et al. (2017a) and a sub-class of Pyne et al. (2017). Otherwise, the Lane et al. (2017) stream class was retained.

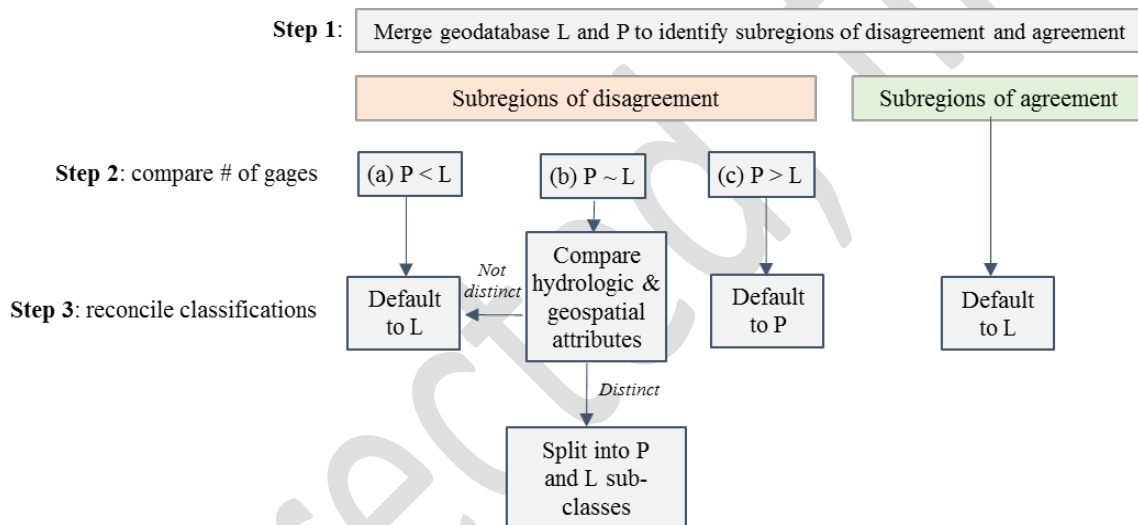


Figure 1. Proposed workflow for reconciling the regional hydrologic classifications by Lane et al. (2017) (L) and Pyne et al. (2017) (P).

Statistical distinctions between the resulting stream classes were evaluated using the Wilcoxon signed-rank test (Wilcoxon 1945) and the matched-paired sign test (Helsel and Hirsch 1992). The un-paired signed-rank statistical test evaluates the overall population of flow data as without consideration for the timing or daily variations. Alternatively, the non-parametric matched-paired test can be used to compare paired time series data regardless of their probability density function.

2.2. Traditional hydrologic baseline characterization

Thirty-four common flow metrics were calculated, representing a non-redundant ecologically-significant distillation of a much larger set of metrics that includes daily, monthly, and annual variability metrics (Konrad et al. 2008). These metrics were calculated for each

reference gauge, and results were evaluated across stream classes. This analysis was intended to illustrate a common use of flow metrics which is to select a pre-determined set of metrics considered in the literature to reflect the attributes described above.

2.3. Dimensionless reference hydrographs (DRHs)

2.3.1. Stream class DRH generation

While archetypes as a whole are useful, an important challenge is to define archetypes that are not specific to a single site; hence the notion of a DRH that is representative for many different stream reaches across a region that receive different absolute amounts of rainfall. To develop DRHs for each stream class, daily streamflow time series from available reference gauges were first non-dimensionalized by dividing each daily flow value by the average daily flow for that water year. This step was performed for each Julian date of each year over a 20-year period from 1968-1988 or 1989-2009, depending on data availability. This period of record was selected to be consistent with the analysis by Lane et al. (2017). Next, the 10th, 25th, 50th, 75th, and 90th percentile dimensionless flows were calculated for each Julian date across the 20-year period. This process was completed for each reference gauge, and then the same percentiles were calculated across all gauge stations in a stream class. For example, the 25th percentile dimensionless flow for October 1 was calculated across 20 years of data for all 25 reference gauges in the snowmelt stream class for a total sample size of 500 for that Julian date in that stream class. The 50th percentile stream class DRH generated through this workflow is then the median hydrologic baseline archetype.

It should be noted that limitations in reference streamflow data may have a major impact on the characterization of natural baselines. In California, there has been over 200 years of land use change as well as natural multi-decadal periods of wetter and drier conditions (Guinn 1890). In other parts of the world, human alterations to flow regimes could span thousands of years. Thus, these recent periods likely do not reflect the full range of natural hydrologic variability, but they represent the best data we have and enable methodological developments. A previous trend analysis of climate non-stationarity (Kendall 1975) and autocorrelation (Durbin and Watson 1950) in the streamflow records from 1968-1988 supported the use of selected streamflow records (Lane et al. 2017).

2.3.2. DRH based hydrologic regime characterization

The stream class DRHs were then used to characterize the reconciled stream classes based on aspects of inter-annual, seasonal, and daily flow variability expected to augment information obtained from traditional flow metrics. Specifically, the DRHs characterized the following hydrologic baseline attributes for nine regional stream classes (Figure 2): (i) long-term average daily flow patterns, (ii) long-term average patterns of inter-annual variability, (iii) seasonal patterns of daily and inter-annual variability, (iv) daily patterns of variability, and (v) seasonal

timing patterns. This set of analyses is not meant to be comprehensive, but rather to highlight a few key aspects of hydrologic baselines that may not be well represented by flow metrics.

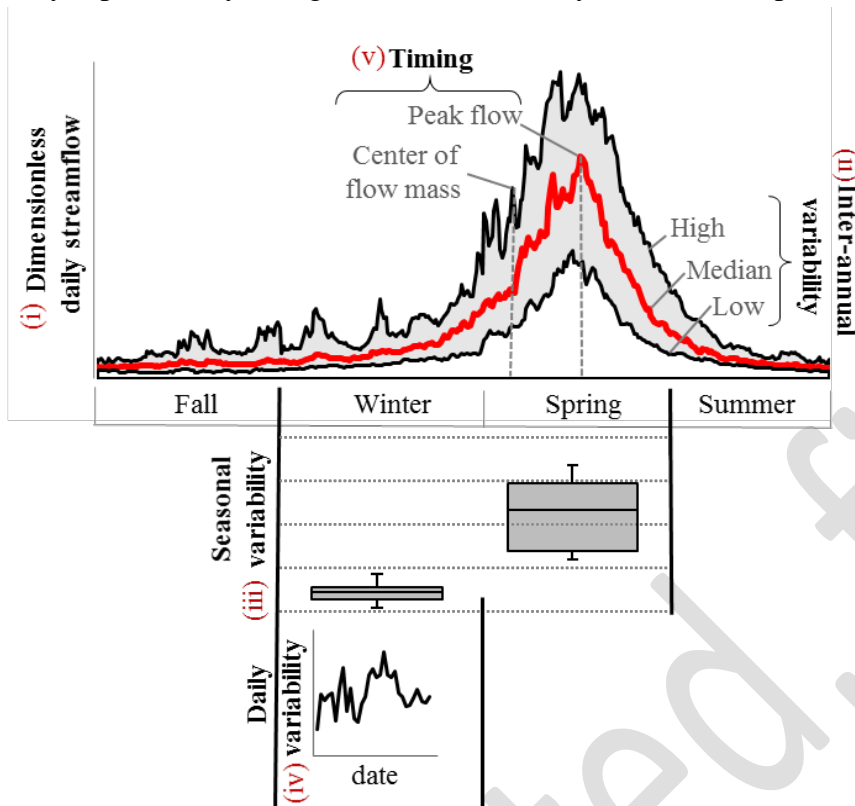


Figure 2. Archetypal hydrologic baseline attributes extracted from stream class DRHs: (i) long-term average daily streamflow patterns, (ii) long-term average patterns of inter-annual variability, (iii) seasonal patterns of daily and inter-annual variability, (iv) daily patterns of variability, and (v) timing of key hydrologic conditions.

(i) Long-term average daily flow patterns

Pairwise comparisons of the median stream class DRH time series (red line in Figure 2) based on R^2 values allowed for a simultaneous comparison of temporal and spatial variability patterns within and between stream classes. The median stream class DRH value for a single Julian date represents the median dimensionless daily flow (as a multiple of average annual flow) across all gage - years in that stream class.

(ii) Long-term average patterns of inter-annual variability

Pairwise stream class comparisons of the 10th, 50th and 90th percentile DRH time series based on R^2 values were then used to compare stream classes across low, average, and high flows, respectively, for the eight most highly correlated stream class pairs based on the previous analysis (i).

(iii) Seasonal patterns of daily and inter-annual variability

The interquartile range (75th - 25th percentile, IQR) of daily streamflow values was evaluated for each stream class both over a water year (Oct 1 - Sept 30) and in each season: fall (Sep 15 - Nov 30), winter (Dec 1 - Feb 28), spring (Mar 1 - Jun 30), and summer (Jul 1 - Sep 15). The stream class IQR for a single Julian date captures the inter-annual variability of flow on that date across all stream class reference gages. It is therefore a dual measure of temporal and spatial variability in daily flows. In each stream class, these daily IQR values were calculated across seasons to assess seasonal differences in daily variability. Finally, these seasonal patterns were compared across stream classes.

(iv) Daily patterns of variability

Daily IQR was then evaluated directly to investigate stream class specific patterns of interannual variability not distinguishable at the annual or seasonal scale.

(v) Seasonal timing patterns

Finally, key seasonal timing attributes were extracted from stream class DRHs, including the date of the center of mass of annual flow and the date of the peak of the wet season. The center of mass date was defined as the median date of peak flow events (defined as periods when the 90th percentile dimensionless daily flow value >4 , representing flows 4 times greater than annual average conditions). The peak wet season date was defined as the date of occurrence of the maximum value of the 50th percentile stream class DRH.

3. Results

3.1. Final reconciled hydrologic classification

3.1.1. Reconciling overlapping classifications

Many subregions of agreement were identified between the two hydrologic classifications, including Snowmelt (SM) and Low-volume Snowmelt and Rain (LSR) in southern California, as named in Lane et al. (2017) (Figure 3). Stream classes with higher agreement were generally elevation-controlled, highlighting elevation as a primary control on streamflow response in California. The strong similarities between the classifications likely reflect the fact that both efforts based their work on first-principles of the interaction of climate, topography, and geology on hydrologic patterns.

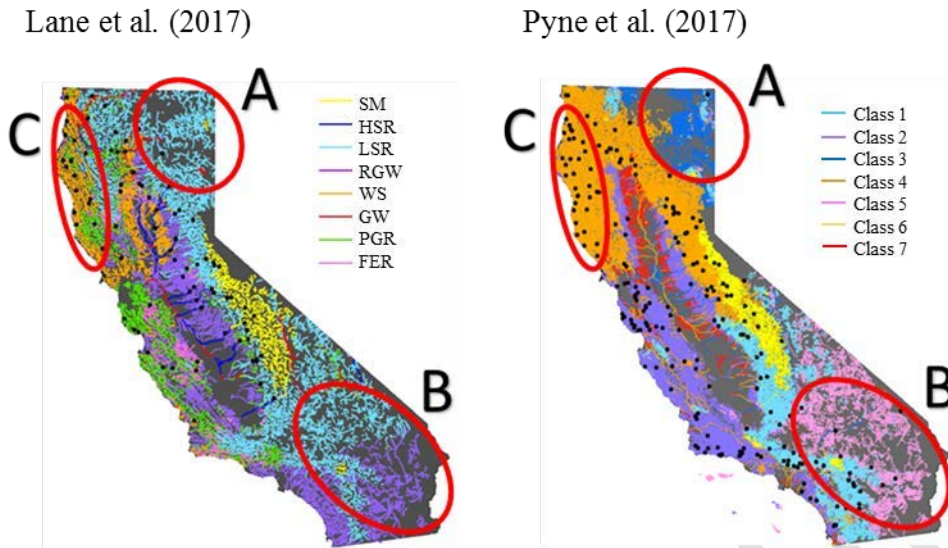


Figure 3. Regions of major differences between the two California hydrologic classifications include the A. Modoc Plateau, B. Mohave, and C. North Coast regions overlaid by reference gauges used for each classification.

Three subregions in particular (circled in Figure 3) exhibited striking differences between the classifications (Figure 3) that were explicitly addressed through the reconciliation workflow (Figure 1): (a) the Modoc Plateau, (b) the Mohave Desert, and (c) the North Coast. The Modoc Plateau subregion in northeastern California was distinguished by Pyne et al. (2017) for its unique combination of high elevation and low precipitation. Because Pyne et al. (2017) had more reference gages in this region (Figure 3), the Pyne et al. (2017) Modoc stream class was selected as the default in that area and incorporated into the final classification (Figure 3, outcome c). Similarly, all Lane et al. (2017) Rain and seasonal Groundwater (RGW) reaches within the Mohave (Figure 3) were changed to Pyne et al. (2017) Flashy Ephemeral Rain (FER) class in areas that overlapped with Pyne et al. (2017)'s Class 5 (Figure 3, outcome c). The FER class was applied here rather than creating a new stream class because FER and Class 5 exhibit very similar hydrologic conditions, and FER stream reaches already exist in other locations in the state. Lane et al. (2017) distinguished two stream classes for the North Coast (Figure 3), the low-volume snowmelt and rain (LSR) and winter storm driven (WS) northwest coast region, while Pyne et al. (2017) only distinguished one class very similar to WS. In this case, since the number of reference gages in this subregion was similar across classifications, the hydrologic and physical attributes of the conflicting stream classes were then evaluated to determine if the physical basis and/or statistical patterns were different enough to warrant splitting (Figure 3, outcome b). The evident differences in the seasonal hydrologic patterns and dominant water sources were used to justify the split between LSR and WS. Specifically, LSR streams exhibited far greater snowmelt influence, including a predictable snowmelt recession in late spring, while WS streams had their center of mass of flow in winter with highly unpredictable winter storm driven high flows dominating the hydrograph.

3.1.2. Final unified hydrologic classification

The reconciliation process generated a consensus classification consisting of nine stream classes, including classes from each of the two independent efforts (Figure 4). These stream classes represent distinct hydrologic landscapes, with distinct flow patterns, flow sources, hydrologic characteristics, and catchment controls over rainfall-runoff response (Wigington 2012). Results include an attributed NHD-plus stream network of the State of California (Figure 4), the stream length distribution by class, and a table summarizing major hydrologic and geospatial variables identified as significant for each stream class (see *Supplemental Materials*).

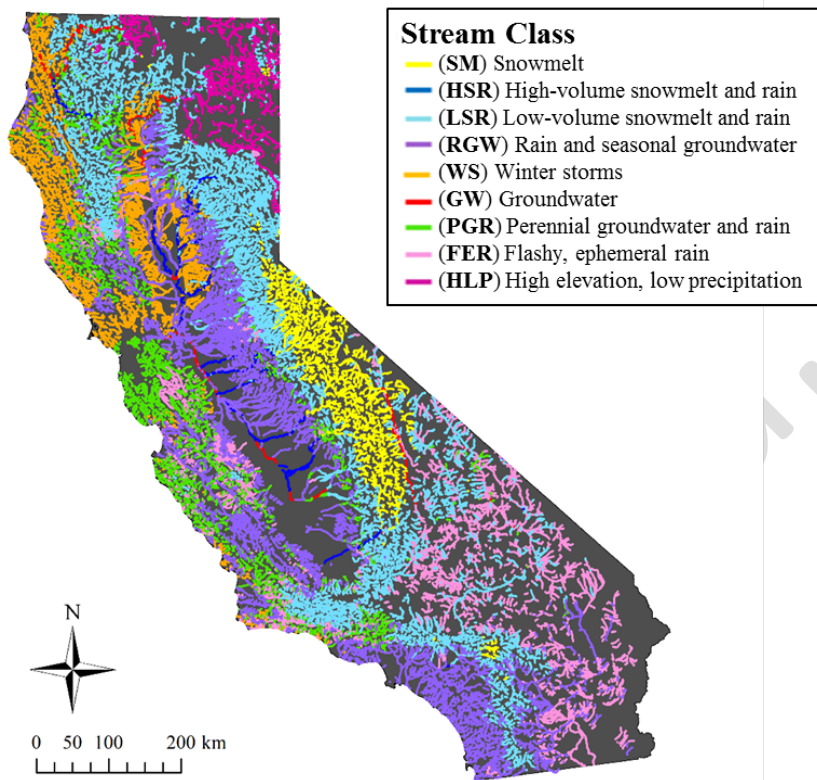


Figure 4. Final reconciled hydrologic classification for the State of California distinguishing nine stream classes.

The matched-paired sign test (Helsel and Hirsch 1992) based on pairwise comparisons of 50th percentile stream class DRH time-series indicated that each stream class was statistically distinct ($\alpha < 0.05$) (Table 1). In contrast, the un-paired Wilcoxon rank-sum test (Wilcoxon 1945) only identified a subset of stream classes (86%) as being statistically distinct whereas other paired stream classes were not (e.g. SM-LSR, RGW-WS, RGW-PGR, WS-PGR and WS-FER). The unpaired rank-sum test, like many flow metrics, evaluates the overall population of streamflow data as a whole without considering daily timing. The fact that the paired statistical test was able to distinguish statistically significant hydrologic differences that the un-paired test was not is a testament to the importance of daily scale variability patterns often not captured by long-term averaged flow metrics.

Table 1. P-values from the matched-paired sign test (**bold**) indicate that all stream classes are statistically distinct ($\alpha < 0.05$), while results from the un-paired Wilcoxon rank-sum test (*italics*) identify fewer stream class distinctions (shaded cells indicate p-values that are not statistically significant). These results demonstrate the importance of daily streamflow time series for characterizing significant aspects of hydrologic variability.

HSR	5.0E-07 <i>1.8E-02</i>							
LSR	3.6E-02 <i>9.5E-01</i>	8.9E-03 <i>2.8E-02</i>						
RGW	3.4E-10 <i>1.6E-11</i>	4.9E-12 <i>1.5E-07</i>	3.3E-62 <i>3.9E-13</i>					
WS	2.8E-05 <i>4.6E-08</i>	2.5E-06 <i>7.4E-07</i>	7.6E-16 <i>2.9E-10</i>	2.9E-08 <i>4.2E-01</i>				
GW	0.0E+00 <i>1.0E-35</i>	0.0E+00 <i>2.8E-36</i>	0.0E+00 <i>7.9E-33</i>	6.2E-81 <i>0.0E+00</i>	0.0E+00 <i>4.4E-41</i>			
PGR	9.3E-08 <i>3.7E-12</i>	6.7E-06 <i>4.1E-09</i>	4.2E-28 <i>2.1E-13</i>	2.8E-02 <i>5.3E-01</i>	4.9E-12 <i>2.5E-01</i>	4.6E-80 <i>0.0E+00</i>		
FER	2.3E-13 <i>2.2E-16</i>	4.2E-11 <i>3.3E-13</i>	3.7E-30 <i>0.0E+00</i>	8.5E-09 <i>6.1E-03</i>	1.6E-04 <i>9.1E-01</i>	3.3E-79 <i>0.0E+00</i>	2.7E-23 <i>1.6E-02</i>	
HLP	0.0E+00 <i>1.1E-18</i>	0.0E+00 <i>6.2E-17</i>	3.4E-10 <i>1.3E-08</i>	0.0E+00 <i>2.1E-28</i>	4.6E-09 <i>2.9E-16</i>	3.2E-24 <i>0.0E+00</i>	2.2E-16 <i>3.0E-27</i>	0.0E+00 <i>2.1E-28</i>
	SM	HSR	LSR	RGW	WS	GW	PGR	FER

3.2. Traditional hydrologic baseline characterization

An evaluation of stream class hydrology based on 34 common flow metrics (from Konrad et al. 2008) yielded some general differences between stream classes (Figure 5 and *Supplemental Materials* Figure 4 and Table 3). For example, snowmelt streams exhibited the most predictable flow patterns overall as indicated by tight clustering of gages (e.g., MaxMonth, MinMonth, RBI, HighDur), lowest flashiness (RBI), few high flow events (HighNum), and shorter low flow durations (LowDur) (Figure 5). In contrast, flashy ephemeral streams exhibited unpredictable conditions, particularly in the duration of low flows (lowDur), the highest (MaxMonth) and lowest (MinMonth) flow months, and hydroperiod. They also exhibited high daily flow variability (RBI, PDC, BFR), many no (MediandNoFlowDays) and low (Per_LowFlow) flow days, and lower average flows (Qmean, Qmed) than other stream classes. Stream class RBI comparisons in particular distinguished broad-scale differences in flashiness, ranging from an average of 0.06 in High elevation Low Precipitation (HLP) streams to 0.55 in Flashy Ephemeral Rain (FER) streams. By contrast, Low-volume Snowmelt and Rain (LSR) and Winter Storm (WS) streams exhibited highly overlapping flow metric values. While WS streams generally exhibited less spread than LSR streams, only MinMonth and HighDur clearly distinguished WS streams due to the much tighter clustering. It should also be noted that for two classes,

Groundwater (GW) and High elevation Low Precipitation (HLP), the low number of reference gauges made it difficult to discern any clear patterns.

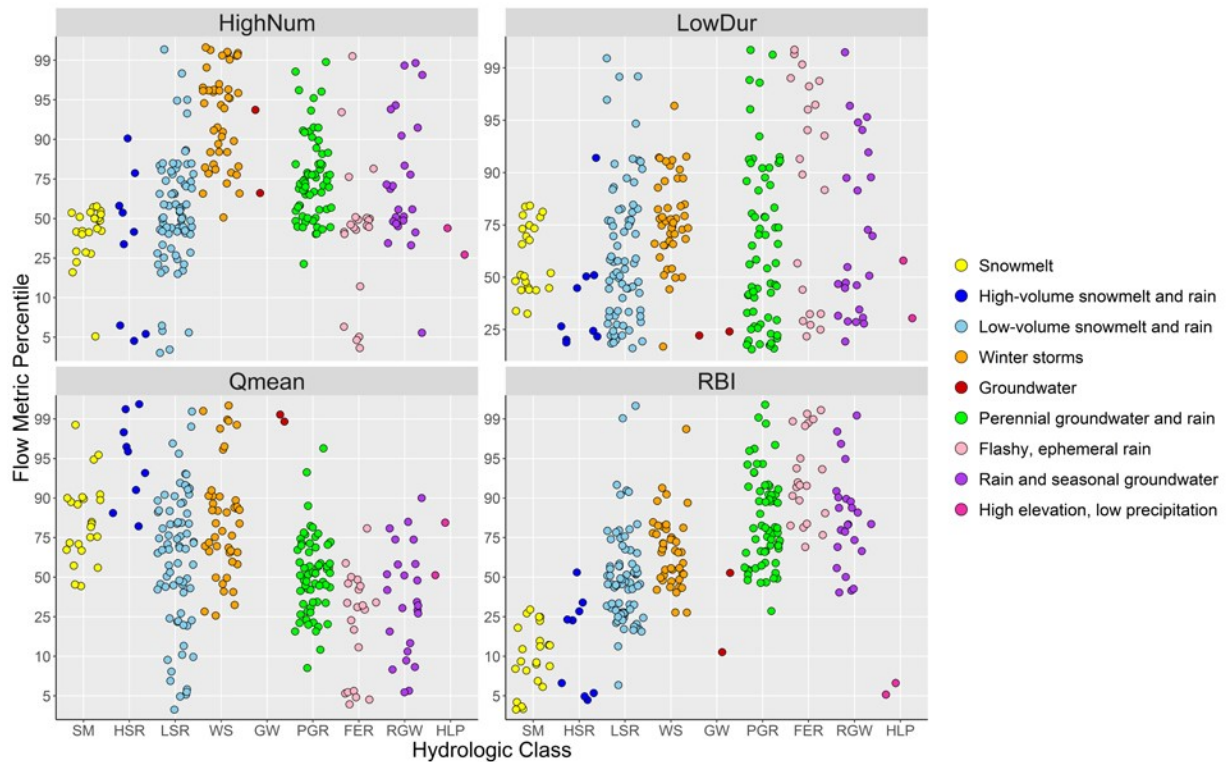


Figure 5. Four common flow metrics (from a set of 32 flow metrics from Konrad et al. (2008)) calculated for each reference gage in each stream class including: number of high flow events (HighNum), low flow duration (LowDur), average annual flow (Qmean) and flashiness (RBI).

While broad differences were evident using the 34 flow metrics alone, limited information content and substantial spread made interpretation difficult for all but the most distinctive stream classes and metrics. While other metrics not considered here may have better captured the patterns of variability that happened to be most distinctive among these stream classes, no comparison of variability patterns across time-scales was possible because metrics were all calculated over some population of streamflow data (e.g., period of record, monthly) rather than retaining daily timing attributes. We expect that combining flow metric assessment as described above with continuous hydrologic baseline archetypes provided by the stream class DRHs will provide more context for these metric values and result in a more complete understanding of the flow patterns than any limited set of metrics alone.

3.3.1. Final DRHs of reconciled stream classes

Dimensionless hydrologic baseline archetypes, or stream class DRHs, simultaneously illustrated inter- and intra-annual streamflow patterns at a daily resolution. Dimensionless daily streamflow, representing the number of times greater any daily flow value is than average annual flow in that water year, varied substantially across seasons and stream classes (Fig. 6). For

example, the 90th percentile DRH (upper dark blue lines in Figure 6) was five times average annual flow in Snowmelt (SM) but surpassed 18 times average annual flow in Winter Storms (WS). These results indicate that inter-annual flow variability can be much greater in WS than SM reaches. The seasonal patterns also differed substantially between these two classes. Across SM streams, as illustrated by the stream class DRH, the vast majority of flow occurs over a short period in late spring (75% of annual streamflow Mar - Jul), while in the WS streams high flows occur over an extended duration during winter (68% of annual streamflow Dec – Mar) (Fig. 6).

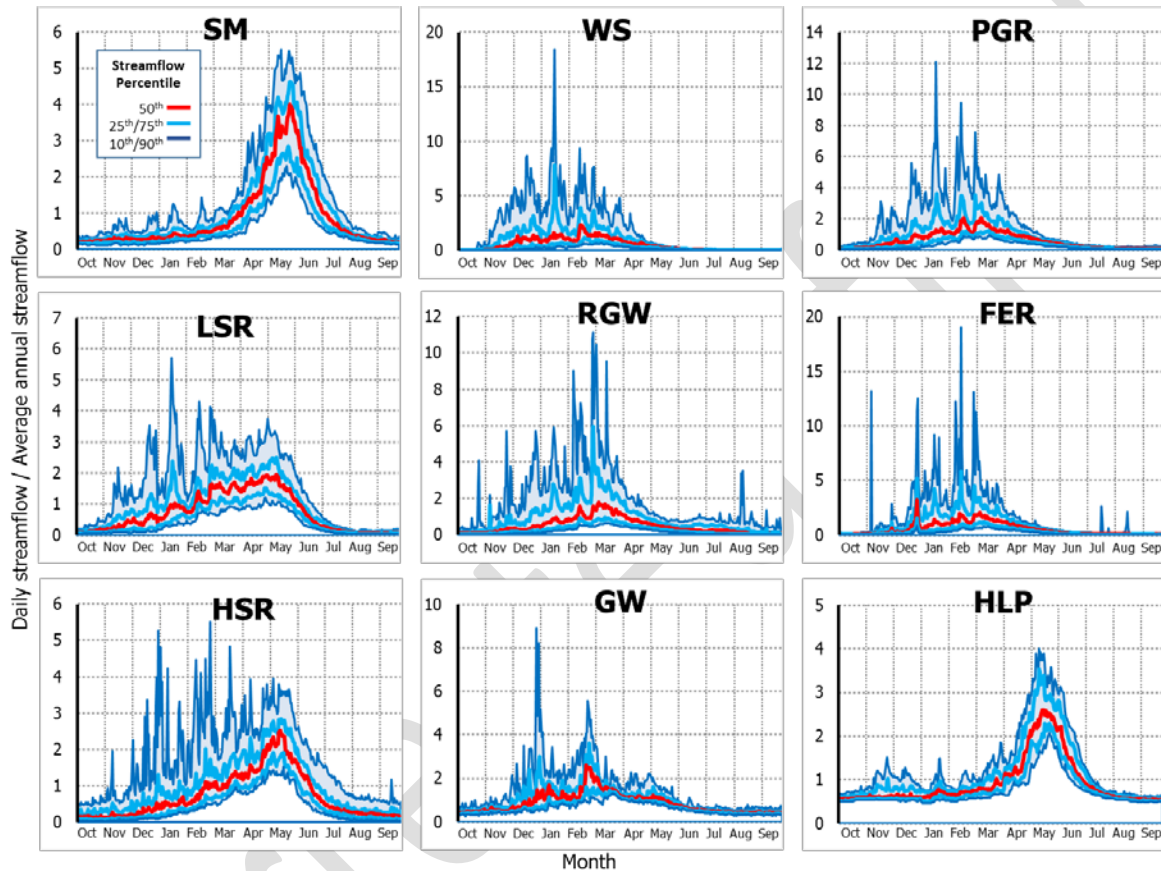


Figure 6. Dimensionless reference hydrographs (DRHs) of the nine reconciled stream classes identified in California: Snowmelt (SM), Low-volume snowmelt and rain (LSR), High-volume snowmelt and rain (HSR), Winter storms (WS), Rain and seasonal groundwater (RGW), Groundwater (GW), Perennial groundwater and rain (PGR), Flashy ephemeral rain (FER), High elevation low precipitation (HLP).

These hydrologic baseline archetypes promote a process-driven understanding of these stream classes that augments the information gleaned from flow metrics. Winter Storm streams are shown to exhibit unpredictable winter high flows driven by winter rain storms and extended extreme low flows in the summer [average median September streamflow 33 cfs; average base flow index 0.01], while Snowmelt streams exhibit highly predictable spring snowmelt recession patterns (i.e., low inter-annual variability and short timing window) (Lane et al. 2017). The Lowvolume Snowmelt and Rain DRH illustrates the combined signature of both the Snowmelt and Winter Storm archetypes, with an evident snowmelt recession associated with limited flow

variance (10th to 90th percentile flow difference <1) as well as a large, extended mass of flow with higher variability (up to 5 time average annual flow) during the winter (Figure 6).

3.3.2. DRH based hydrologic regime characterization

Statistical analysis of stream class DRHs revealed distinctive aspects of hydrologic baselines including: long-term average daily streamflow patterns, long-term average patterns of interannual variability, seasonal patterns of inter-annual variability, daily patterns of variability, and seasonal timing patterns.

(i) Long-term average daily flow patterns

The pairwise comparisons of median stream class DRHs exhibited a range of R² values ranging from 0 to 0.95, indicating that some stream classes were much more similar than others (Table 2). Snowmelt and High-elevation Low Precipitation were the most similar stream classes based on this index (R²=0.95), followed by Perennial Groundwater and Rain (PGR) and Flashy Ephemeral Rain (FER) (R²=0.80) and PGR and Groundwater (GW) (R²=0.78). Interestingly, FER and GW were far less similar (R²=0.66), confirming that PGR represents a combination of end-member patterns.

Table 2. R² values of pairwise stream class comparisons for 50th percentile DRH time series. Bolded values indicate eight most highly correlated stream class pairs. Stream classes are as follows: Snowmelt (SM), Low-volume snowmelt and rain (LSR), High-volume snowmelt and rain (HSR), Winter storms (WS), Rain and seasonal groundwater (RGW), Groundwater (GW), Perennial groundwater and rain (PGR), Flashy ephemeral rain (FER), High elevation low precipitation (HLP).

SM									
HSR	0.68								
LSR	0.26	0.71							
RGW	0.03	0.04	0.40						
WS	0.02	0.27	0.64	0.59					
GW	0.00	0.12	0.54	0.61	0.59				
PGR	0.00	0.13	0.58	0.72	0.70	0.78			
FER	0.03	0.33	0.76	0.55	0.74	0.66	0.80		
HLP	0.95	0.70	0.27	0.02	0.02	0.00	0.00	0.02	
	SM	HSR	LSR	RGW	WS	GW	PGR	FER	HLP

(ii) Long-term average patterns of inter-annual variability

Further comparison of low and high flow patterns (the 10th and 90th percentile DRHs, respectively) for the eight most highly correlated stream class pairs (based on the 50th percentile DRHs, Table 2) provides additional insight into stream class-wise differences in inter-annual variability (Table 3). For example, Winter Storm (WS) and FER classes exhibited similar average streamflow patterns (R²=0.74) and low flow patterns (R²=0.91), but dissimilar high flow

patterns ($R^2=0.50$). Thus, while the two stream classes have similar seasonal trends, reflecting winter rain storms as the primary flow source, they can be distinguished by contrasting high flow patterns that occur in wet years. The WS 90th percentile DRH tracks the median daily DRH and has a maximum of 13, while the FER 90th percentile DRH diverges substantially from the median and has a maximum of 40. These results indicate that FER streams have much flashier, unpredictable wet year high flows but similar overall patterns of seasonality compared to WS streams. Alternatively, Snowmelt and High-elevation Low Precipitation streams exhibited similar patterns across low, average and high flows, indicating similar hydrologic responses in all water year types.

Table 3. R^2 values of select pairwise stream class comparisons of low (10th percentile), average (50th percentile), and high (90th percentile) flows. Stream classes are as follows: Snowmelt (SM), Low-volume snowmelt and rain (LSR), High-volume snowmelt and rain (HSR), Winter storms (WS), Rain and seasonal groundwater (RGW), Groundwater (GW), Perennial groundwater and rain (PGR), Flashy ephemeral rain (FER), High elevation low precipitation (HLP).

	R^2		
	0.1	0.5	0.9
SM-HLP	0.92	0.95	0.90
PGR-FER	0.83	0.80	0.69
GW-PGR	0.66	0.78	0.78
LSR-FER	0.81	0.76	0.73
WS-FER	0.91	0.74	0.50
RGW-PGR	0.73	0.72	0.51
LSR-HSR	0.78	0.71	0.41
WS-PGR	0.80	0.70	0.45

In general, high flow patterns were most distinct across stream classes while low flow patterns were most similar (Table 3). For example, Low-volume and High-volume Snowmelt and Rain stream classes exhibited very similar low and average flow DRHs ($R^2=0.78$ and 0.71 , respectively), as they are both snow-and-rain driven, but they had distinctive high flow patterns ($R^2=0.41$) (Table 3) due to flashy and high magnitude winter rain-dominated storms in Low volume Snowmelt and Rain streams. This finding indicates that wet years provide unique hydrologic patterns in different stream classes that should be considered in environmental flows setting. An exception to this trend was Groundwater and Perennial Groundwater and Rain stream classes, which illustrated the opposite pattern: low flow patterns were contrasting (specifically, the stability of baseflows over a year), while average and high flows were similarly infrequent and occurred primarily in winter.

(iii) Seasonal patterns of inter-annual variability

The inter-quartile range (IQR) was calculated at a daily time-step and can therefore be used to compare annual and seasonal patterns of inter-annual flow variability across stream classes. In a comparison of long-term inter-annual variability at the annual scale (Figure 7, top plot), Snowmelt (SM) exhibited the highest median daily IQR (0.49), followed by Low-volume Snowmelt and Rain (LSR) (0.43), while Groundwater (GW) had the lowest value (0.06). These

trends differed significantly between seasons (Figure 7, four bottom plots). For example, SM median daily IQR ranged from 0.15 in Fall to 1.15 in Spring, within a single stream class. Results also demonstrate that different stream class DRHs have more or less inter-annual variability in different seasons. Winter Storms (WS) streams exhibited the highest median daily IQR in Fall, Rain and Seasonal Groundwater (RGW) streams in Winter, and Snowmelt streams in Spring and Summer.

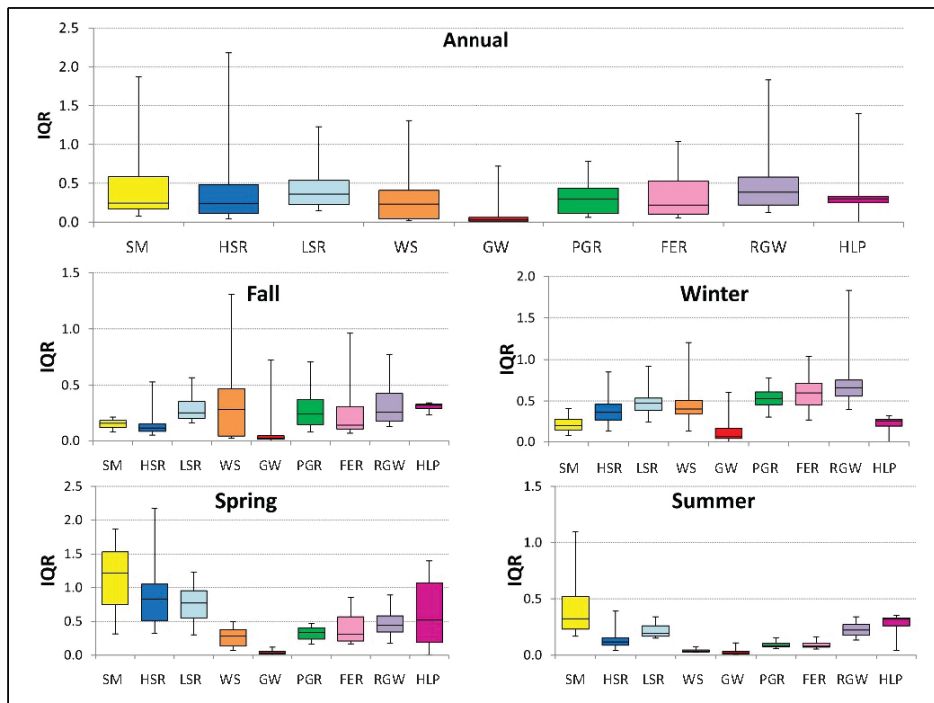


Figure 7. Box-plots illustrate the interquartile range (IQR) of daily flow across stream classes over the period of record both annually and seasonally for fall, winter, spring, and summer. Stream classes are as follows: Snowmelt (SM), Low-volume snowmelt and rain (LSR), High-volume snowmelt and rain (HSR), Winter storms (WS), Rain and seasonal groundwater (RGW), Groundwater (GW), Perennial groundwater and rain (PGR), Flashy ephemeral rain (FER), High elevation low precipitation (HLP).

A selected comparison of daily IQRs between winter and spring highlights major seasonal differences in variability patterns between stream classes (Figure 8). In winter, median daily IQR was significantly higher in Winter Storm (WS) than Snowmelt (SM) streams while the reverse is true in spring. In winter, this distinction is likely driven by the high variability in timing, magnitude, and duration of winter rain storms contrasted by the predictable winter low flows associated with high elevation snowmelt-dominated streams. In spring, by contrast, inter-annual and spatial differences in climate patterns drive high daily variability in the snowmelt recession in Snowmelt streams, while high variability rain storms are less common by that time of year. The Low-volume Snowmelt and Rain (LSR) class exhibited a balance of the physical drivers of both SM and WS classes, resulting in intermediate values in both seasons. However, the variable snowmelt influence of LSR streams drove a higher median IQR in spring than winter. Finally, while WS streams had similar median IQR values in winter and spring, the range of daily IQRs

exhibited (indicated by the boxplot whiskers in Figure 8) was much greater in winter than spring, indicating a need to evaluate daily patterns of variability that are not discernable at the seasonal scale.

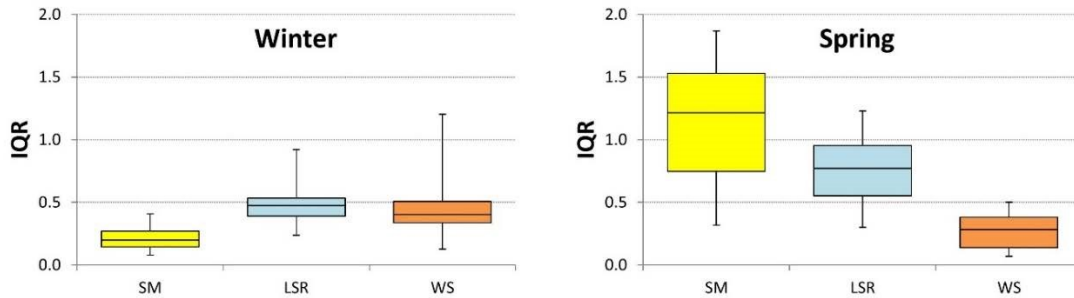


Figure 8. Boxplots summarizing seasonal patterns of daily IQR of selected stream classes [Snowmelt (SM), Low-volume Snowmelt and Rain (LSR), Winter Storms (WS)] in winter and spring highlight seasonal and sub-seasonal differences in interannual streamflow variability.

(iv) Daily patterns of flow variability

While the median daily IQR in spring was similar in WS and FER stream classes, daily IQR patterns varied significantly. For instance, over the month of March, WS daily IQR remained relatively stable, fluctuating by only 0.11 (Figure 9). By contrast, FER daily IQR varied by over 0.5 during the same period and exhibited a nearly inverted pattern of increasing and decreasing daily IQR with WS.



Figure 9. Comparison of two stream classes [Winter Storms (WS) and Flashy, Ephemeral Rain (FER)] with similar seasonal average daily IQRs highlights additional patterns of flow variability at the daily scale.

(v) Seasonal timing patterns

Finally, by querying timing attributes of low, median and high flows (i.e. 10th, 50th, or 90th percentile) obtained directly from the DRHs, key differences in timing between stream classes were revealed with minimal calculation requirements (Table 4). For example, Snowmelt streams were shown to have very similar timing for the center of mass of flow (May 26) and the peak wet season flow (May 31). By contrast, High-volume Snowmelt and Rain streams exhibited a center

of mass in winter (Feb 12) and a wet season peak lagged by over two months (May 15), emphasizing the dual controls of rain and snowmelt on the hydrograph.

Table 4. Timing attributes obtained from stream class DRHs, including the median date of peak flow events (defined as days when the 90th percentile DRH > 4) and the date of the maximum of the 50th percentile DRH. These timing attributes represent the center of mass of annual flow and the peak of the wet season, respectively.

	SM	HSR	LSR	WS	GW	PGR	FER	RGW	HLP
median peak date [90th]	26-May	12-Feb	3-Mar	24-Jan	4-Jan	14-Feb	12-Feb	19-Feb	15-May
date of max of 50th	31-May	18-May	15-May	17-Jan	23-Feb	28-Dec	27-Dec	9-Mar	18-May

4. Discussion

4.1. Reconciling regional hydrologic classifications

The reconciled classification drew from the strengths of the two previous classifications, which were each developed for slightly different purposes. Class differentiation in areas with low gauge density was generally greater based on the catchment derived clustering used by Pyne et al. (2017), whereas the stream gauge based classification used by Lane et al. (2017a) was able to better separate streams in similar physical setting, but with subtle (yet important) hydrologic differences. This demonstrates the value of the integrative and interactive classification developed through the reconciliation process.

Traditional flow metrics showed expected patterns across stream classes suggesting good hydrologic differentiation using the reconciled classification. For example, systems characterized by snowmelt and groundwater were generally less variable than those characterized by winter rains and flashy conditions. As expected, flow metrics for stream classes with more reference gauges covered larger ranges of values and more overlap with similar hydrologic categories. Such “boundary effects” are common in most classification systems and suggest that a secondary classification, such as one based on geomorphology or hydraulic properties, may be useful in further differentiating streams into subclasses that have more utility for environmental flow analysis and restoration planning.

4.2. The role of archetypes in ecohydrology

Archetypes play a vital role in science because they enable experimental testing to isolate the defining conditions and dynamics of clearly differentiated entities. In geomorphology, a landform archetype ideally reflects a unique process set that creates it, and in turn serves as a driver of ecohydrological functions (Lane et al. 2018) as well as subsequent morphodynamics (Brown et al. 2015). The same is true for a biological archetype involving a specific community structure - it is formed by a unique suite of processes and in turn drives other conditions as well as feedbacks (Deutschman et al. 1997; Parasiewicz et al. 2008). This study demonstrates how hydrologic baselines archetypes, termed dimensionless reference hydrographs (DRHs), can be produced to not only reflect and provide context for lumped statistical flow metrics, but characterize multi-scale continuous daily, seasonal, and inter-annual flow patterns. These DRHs

provide a link between flow metrics and environmental management objectives by providing additional insight on hydrologic drivers of ecological response. Addressing study objective 5, the following section discusses the physical and ecological interpretability of the hydrologic archetypes compared to traditional flow metrics.

4.2.1. Ecological utility of hydrologic archetypes

The importance of flow variability and cyclical patterns of intra- and inter-annual variation for shaping the biophysical attributes and functioning of river systems is well recognized (Naiman et al. 2008; Poff et al. 1997). While traditional flow metrics are generally suitable for characterizing *major differences in long-term average conditions* (including flow variability and predictability), study results indicated that DRHs may be more suitable for describing *multi-scale patterns of variability*, as evaluated at the daily, seasonal, and inter-annual scale.

While all calculated flow metrics for Low- and High-volume Snowmelt and Rain stream classes exhibited overlapping values (Figure 5; Supplemental Figure 4), ecologically distinct flow patterns were revealed when combined with a direct comparison of their stream class DRHs, and thus placed into the context of daily and seasonal variability. Both stream classes were influenced by winter rain storms, exhibiting flashier winter flows than Snowmelt streams (RBI: HSR, LSR, SM-) and winter maximum flows (MaxMonth: HSR, LSR, SM-). However, the DRHs (Figure 6) and seasonal IQR analysis (Figure 7) demonstrated that HSR streams exhibited more, larger winter storm events than LSR streams (Figure 6). These storms mobilize large amounts of sediment that are re-sorted and distributed during the spring snowmelt recession (Yarnell et al. 2015). The DRHs also highlighted increased spring streamflow variability in HSR than LSR (Figure 6), reflecting a larger snowmelt influence from upstream catchments included in the larger drainage areas of the High-volume stream class. The Highvolume DRH also exhibited an extended spring recession, with flows remaining greater than average annual flow until mid-July in wet years (Figure 6). This longer duration spring recession provides ample opportunity for native species to reproduce and migrate (Yarnell et al. 2010). Thus HSR streams, with greater potential sediment mobilization and longer duration spring snowmelt recessions, may be more likely to provide habitat conditions conducive to spring spawning and migration as well as cooler water temperatures in summer due to a shortened low flow period.

Stream class DRH comparisons also offer insight into the ecological mechanisms and life history strategies contributing to habitat suitability for native species. For example, the high daily and inter-annual flow variability exhibited by Winter Storms and Flashy Ephemeral Rain streams in winter (Figure 7) indicate flow regimes with low predictability for resident aquatic species. Availability of aquatic refugia from unpredictable high magnitude winter storms, such as connections to backwater or floodplain habitat favored by juvenile salmon (Moyle 2002), may be of much greater significance in these streams than in the more predictable streams. In contrast, the high magnitude (Figure 5, Figure 6) and low daily variability (Figure 7) of groundwater influenced streams (e.g., GW, PGR) in summer may provide refuge conditions for native coldwater species (e.g. *Oncorhynchus* spp.) that rely on cool perennial water year-round.

Similarly, the stream class DRHs indicated that the timing of the transition from spring snowmelt to summer baseflow in snowmelt-dominated stream classes (e.g., SM, HSR, LSR, HLP) was typically in July, while the start of the low flow season in rain-dominated streams (e.g., WS, FER, PGR, RGW) was in May (Figure 6). For species whose life history strategies rely on the availability of certain spring habitat conditions (e.g. *Populus* spp., Stella et al. 2006), this is critical. Regional management actions and conservation strategies would benefit by taking into account these seasonal differences in flow regime timing, magnitude, and variability across stream classes.

4.2.2. Regional analysis using hydrologic archetypes

By providing a complete but dimensionless characterization of streamflow patterns, DRHs have the capacity to retain critical high resolution information while supporting the broad spatial comparisons needed for regional analysis. McManamay et al. (2012) identified within-class variability associated with scale differences among individual streams as a principle challenge in developing coherent regional streams classes. The dimensionless archetypes proposed here eliminate scale effects, addressing the need to simultaneously consider both spatial and temporal variability in regional ecosystem assessments. Consequently, variability in catchment area among reference gages has limited effect on characterization of hydrologic baselines and allows for direct comparison and flow – ecology hypothesis testing across regional stream classes in addition to individual reference gages.

Some stream classes may inherently provide flow conditions more conducive to certain native species requirements based on their distinct patterns of natural variability (e.g. predictability of winter flows, daily variability of summer baseflow). The ability to quantify these differences in hydrologic baseline conditions across heterogeneous landscapes, when combined with information on native species distributions, is expected to offer insight into potential habitat suitability at the regional scale, thus providing resource managers with much needed data for regional conservation or restoration planning. Results from this study indicate an opportunity to rapidly evaluate differences between regional stream classes across water years, hydrograph components (e.g. low vs high flows), and seasons that can be evaluated with respect to specific ecological and geomorphic management objectives.

4.3. Future applications of hydrologic archetypes in ecohydrology

Hydrologic baseline archetypes as described in this study open up new avenues for hydrologic research. Two specific examples include assessing hydrological alteration due to human activities and evaluating the effects of climatic change. As stream class DRHs provide daily-scale reference hydrologic expectations for a stream class, comparisons with observed continuous streamflow data can provide insight into not only the degree but the *pattern* of flow alteration a particular stream may exhibit. Flow alteration has been evaluated in a variety of ways using traditional flow metrics and comparisons with aquatic communities (Webb et al. 2013; Carlisle et al. 2016). Additional insight into the pattern of alteration (e.g., systemic depletion or

augmentation, alteration to seasonal or inter-annual variability patterns, etc.), and thus potential reasons why aquatic communities may or may not be robust, can be obtained from the more detailed and temporally explicit stream class DRHs. When designing flow regimes to improve impaired flow conditions, seasonally relevant flow management targets based on the distribution of daily or seasonal flows within a stream class DRH can be created, providing water resource managers with increased flexibility over static flow targets representing long-term averages. This is particularly useful when balancing multiple ecological endpoints of interest with water supply objectives.

The detailed temporal data provided in stream class DRHs also allows for comparisons with modeled streamflow conditions, making this technique particularly useful for considering climate change impacts in streamflow restoration planning. Most snowmelt-influenced streams across California and the western U.S. will experience an increase in winter flow magnitude and a decrease in snowmelt runoff and timing as winter storm precipitation shifts from snow to rain in a warmer climate (Null et al. 2010; Ficklin et al. 2016). Comparisons between modeled climate scenarios and DRHs can provide details on potential changes to ecologically relevant patterns of variability, such as timing and predictability of snowmelt or variability of summer baseflow, that are more difficult to discern using traditional annually averaged flow metrics. Over time, the stream class DRHs can then be evaluated for evidence of shifting baselines with regard to inter-annual or intra-annual variability. The continuous daily flow regimes provided by the DRHs can also aid in assessing uncertain future flow conditions associated with climate change and future water resource management decisions by providing reference ranges of variability that can be compared to potential ranges of future flows.

5. Conclusions

This study evaluated the use of daily resolution dimensionless archetypes of regional stream classes derived from a hydrologic classification to support the interpretation of flow metrics and inform regional environmental flows assessments. A novel approach for reconciling existing classification systems resulted in a single robust hydrologic classification to simplify and improve regional environmental flows efforts in California. Hydrologic baseline archetypes (i.e., dimensionless reference hydrographs) generated for each of the nine stream classes were shown to characterize continuous inter-annual, seasonal, and daily streamflow patterns, allowing for explicit, simultaneous evaluation of within and between classes variability in space and time. These high-resolution archetypes provided context for information gleaned from a set of common flow metrics, thus improving understanding of hydrologic drivers and ecological mechanisms contributing to flow variability and habitat suitability. Results indicate that hydrologic baseline archetypes may offer a distinct and useful tool for moving beyond metrics to improve linkages between aquatic patterns and processes and their hydrologic controls and drivers.

Citations

- Archfield, Stacey A, Richard M Vogel, and Sara L Brandt. 2007. "Estimation of flow-duration curves at ungaged sites in southern New England." In *World Environmental and Water Resources Congress 2007: Restoring Our Natural Habitat*, 1-14.
- Arthington, Angela H. 2006. 'The challenge of providing environmental flow rules to sustain river ecosystems', *Ecological Applications*.
- Asquith, William H, Meghan C Roussel, and Joseph Vrabel. 2006. "Statewide analysis of the drainage-area ratio method for 34 streamflow percentile ranges in Texas." In.: US Geological Survey.
- Beechie, Timothy J., David A. Sear, Julian D. Olden, George R. Pess, John M. Buffington, Hamish Moir, Philip Roni, and Michael M. Pollock. 2010. 'Process-based Principles for Restoring River Ecosystems', *BioScience*, 60: 209-22.
- Bender, Donald L, and John A Roberson. 1961. 'The use of a dimensionless unit hydrograph to derive unit hydrographs for some Pacific Northwest basins', *Journal of Geophysical Research*, 66: 521-28.
- Blöschl, Günter. 2001. 'Scaling in hydrology', *Hydrological Processes*, 15: 709-11.
- Blöschl, Günter, and Murugesu Sivapalan. 1995. 'Scale issues in hydrological modelling: a review', *Hydrological Processes*, 9: 251-90.
- Brown, Rocko A., Gregory B. Pasternack, and T. Lin. 2015. 'The Topographic Design of River Channels for Form-Process Linkages', *Environmental Management*, 57: 929-42.
- Brown, Rocko A., Gregory B. Pasternack, and W. W. Wallender. 2014a. 'Synthetic river valleys: Creating prescribed topography for form–process inquiry and river rehabilitation design', *Geomorphology*, 214: 40-55.
- Brown, Stuart C, Rebecca E Lester, Vincent L Versace, Jonathon Fawcett, and Laurie Laurenson. 2014b. 'Hydrologic landscape regionalisation using deductive classification and random forests', *PLoS One*, 9: e112856.
- Bunn, Stuart E., and Angela H. Arthington. 2002. 'Basic Principles and Ecological Consequences of Altered Flow Regimes for Aquatic Biodiversity', *Environmental Management*, 30: 492507.
- Carlisle, Daren M., S. Mark Nelson, and Jason May. 2016. 'Associations of stream health with altered flow and water temperature in the Sierra Nevada, California', *Ecohydrology*, 9: 93041.
- Cullum, Carola, Gary Brierley, George LW Perry, and Ed TF Witkowski. 2017. 'Landscape archetypes for ecological classification and mapping', *Progress in Physical Geography*, 41: 95-123.
- Deutschman, Douglas H, Simon A Levin, Catherine Devine, and Linda A Buttel. 1997. 'Scaling from trees to forests: analysis of a complex simulation model', *Science*, 277: 1684b-84b.
- Durbin, J, and G Watson. 1950. 'Testing for serial correlation in least squares regression-I', *Biometrika*, 37: 409-28.
- Eng, Ken, Theodore E. Grantham, Daren M. Carlisle, and David M. Wolock. 2017. 'Predictability and selection of hydrologic metrics in riverine ecohydrology', *Freshwater Science*, 0: 00000.
- Farmer, William H, and Richard M Vogel. 2013. 'Performance-weighted methods for estimating monthly streamflow at ungaged sites', *Journal of Hydrology*, 477: 240-50.
- Ficklin, Darren L, Sally L Letsinger, Iris T Stewart, and Edwin P Maurer. 2016. 'Assessing differences in snowmelt-dependent hydrologic projections using CMIP3 and CMIP5 climate forcing data for the western United States', *Hydrology Research*, 47: 483-500.

- Gasith, A., and B. Resh. 1999. 'Streams in Mediterranean Regions: Abiotic Influences and Biotic Responses to Predictable Seasonal Event', *Annual Review of Ecological Systems*.
- Guinn, James Miller. 1890. 'Exceptional years: a history of California floods and drought', *Historical Society of Southern California, Los Angeles (1890)*, 1: 33-39.
- Haines, AT, BL Finlayson, and TA McMahon. 1988. 'A global classification of river regimes', *Applied Geography*, 8: 255-72.
- He, Y, A Bárdossy, and E Zehe. 2011. 'A review of regionalisation for continuous streamflow simulation', *Hydrology and Earth System Sciences*, 15: 3539.
- Helsel, Dennis R, and Robert M Hirsch. 1992. *Statistical methods in water resources* (Elsevier).
- Kendall, Maurice. 1975. *Multivariate analysis* (Charles Griffin: London).
- Kennard, Mark J., Bradley J. Pusey, Julian D. Olden, Stephen J. Mackay, Janet L. Stein, and Nick Marsh. 2010. 'Classification of natural flow regimes in Australia to support environmental flow management', *Freshwater Biology*, 55: 171-93.
- Konrad, CP, AMD Brasher, and JT May. 2008. 'Assessing streamflow characteristics as limiting factors on benthic invertebrate assemblages in streams across the western United States', *Freshwater Biology*, 53: 1983-98.
- Lane, Belize A, Gregory B Pasternack, and Samuel Sandoval-Solis. 2018. 'Integrated analysis of flow, form, and function for river management and design testing', *Ecohydrology*.
- Lane, Belize A., Helen E. Dahlke, Gregory B. Pasternack, and Samuel Sandoval-Solis. 2017. 'Revealing the Diversity of Natural Hydrologic Regimes in California with Relevance for Environmental Flows Applications', *JAWRA Journal of the American Water Resources Association*: n/a-n/a.
- Lane, Belize A., Samuel Sandoval-Solis, and Erik C. Porse. 2014. 'Environmental Flows in a Human-Dominated System: Integrated Water Management Strategies for the Rio Grande/Bravo Basin', *River Research and Applications*, 31: 1053-65.
- Leibowitz, Scott G., Randy L. Comeleo, Parker J. Wigington, Marc H. Weber, Eric A. Sproles, and Keith A. Sawicz. 2016. 'Hydrologic Landscape Characterization for the Pacific Northwest, USA', *JAWRA Journal of the American Water Resources Association*, 52: 47393.
- Lundquist, Jessica D, and Daniel R Cayan. 2002. 'Seasonal and spatial patterns in diurnal cycles in streamflow in the western United States', *Journal of Hydrometeorology*, 3: 591-603.
- Lytle, David A, and N LeRoy Poff. 2004. 'Adaptation to natural flow regimes', *Trends in ecology & evolution*, 19: 94-100.
- McCuen, Richard H. 1989. *Hydrologic analysis and design* (Prentice-Hall Englewood Cliffs, NJ).
- McManamay, RA, DJ Orth, CA Dolloff, and EA Frimpong. 2012. 'Regional frameworks applied to hydrology: can landscape-based frameworks capture the hydrologic variability?', *River Research and Applications*, 28: 1325-39.
- Mierau, Darren W., William J. Trush, Gabriel J. Rossi, Jennifer K. Carah, Matthew O. Clifford, and Jeanette K. Howard. 2017. 'Managing diversions in unregulated streams using a modified percent-of-flow approach', *Freshwater Biology*: n/a-n/a.
- Moyle, Peter B. 2002. *Inland fishes of California* (University of California Press: Berkeley, CA).
- Naiman, Robert J., Joshua J. Latterell, Neil E. Pettit, and Julian D. Olden. 2008. 'Flow variability and the biophysical vitality of river systems', *Comptes Rendus Geoscience*, 340: 629-43.
- Null, Sarah E, Joshua H Viers, and Jeffrey F Mount. 2010. 'Hydrologic response and watershed sensitivity to climate warming in California's Sierra Nevada', *PLoS One*, 5: e9932.

- Olden, Julian D., Mark J. Kennard, and Bradley J. Pusey. 2012. 'A framework for hydrologic classification with a review of methodologies and applications in ecohydrology', *Ecohydrology*, 5: 503-18.
- Olden, Julian D., and N. L. Poff. 2003. 'Redundancy and the choice of hydrologic indices for characterizing streamflow regimes', *River Research and Applications*, 19: 101-21.
- Parajka, J., A. Viglione, M. Rogger, J. L. Salinas, M. Sivapalan, and G. Blöschl. 2013. 'Comparative assessment of predictions in ungauged basins – Part 1: Runoffhydrograph studies', *Hydrology and Earth System Sciences*, 17: 1783-95.
- Parasiewicz, Piotr, John Nestler, N LeRoy Poff, and R Andrew Goodwin. 2008. 'Virtual reference river: a model for scientific discovery and reconciliation', *Ecological management: new research. Nova Science Publishers, Inc., New York*: 189-98.
- Pasternack, Gregory B, and Linda A Hinnov. 2003. 'Hydrometeorological controls on water level in a vegetated Chesapeake Bay tidal freshwater delta', *Estuarine, Coastal and Shelf Science*, 58: 367-87.
- Poff, N. L., and J. David Allan. 1995 'Functional organization of stream fish assemblages in relation to hydrologic variability', *Ecology*.
- Poff, N. Leroy, J. David Allan, Mark B. Bain, James R. Karr, Karen L. Prestegard, Brian D. Richter, Richard E. Sparks, and Julie C. Stromberg. 1997. 'The natural flow regime: a paradigm for river conservation and restoration', *BioScience*, 47: 769-84.
- Poff, N. Leroy, Brian D. Richter, Angela H. Arthington, Stuart E. Bunn, Robert J. Naiman, Eloise Kendy, Mike Acreman, Colin Apse, Brian P. Bledsoe, Mary C. Freeman, James Henriksen, Robert B. Jacobson, Jonathan G. Kennen, David M. Merritt, Jay H. O'Keeffe, Julian D. Olden, Kevin Rogers, Rebecca E. Tharme, and Andrew Warner. 2010. 'The ecological limits of hydrologic alteration (ELOHA): a new framework for developing regional environmental flow standards', *Freshwater Biology*, 55: 147-70.
- Poff, N. Leroy, and Julie K. H. Zimmerman. 2010. 'Ecological responses to altered flow regimes: a literature review to inform the science and management of environmental flows', *Freshwater Biology*, 55: 194-205.
- Postel, Richter. 2004. 'Rivers for life'.
- Pyne, Matthew, D. M. Carlisle, C. P. Konrad, and Eric Stein. 2017. 'Classification of California Streams', *Ecohydrology*.
- Reidy Liermann, C. A., J. D. Olden, Tim Beechie, M. J. Kennard, P. B. Skidmore, Christopher P. Konrad, and H. Imaki. 2012. 'Hydrogeomorphic Classification of Washington State Rivers to Support Emerging Environmental Flow Management Strategies', *River Research and Applications*, 28: 1340-58.
- Richter, Brian D. 1996. 'A Method for Assessing Hydrologic Alteration within Ecosystems', *Conservation Biology*.
- Salas, Jose D. 1993. 'Models for Data Generation in Hydrology: Multivariate Techniques.' in, *Stochastic Hydrology and its Use in Water Resources Systems Simulation and Optimization* (Springer).
- Sanborn, Stephen C., and Brian P. Bledsoe. 2006. 'Predicting streamflow regime metrics for ungauged streams in Colorado, Washington, and Oregon', *Journal of Hydrology*, 325: 241-61.
- Sawicz, K., T. Wagener, M. Sivapalan, P. A. Troch, and G. Carrillo. 2011. 'Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA', *Hydrology and Earth System Sciences*, 15: 2895-911.

- Sivapalan, Murugesu. 2003. 'Prediction in ungauged basins: a grand challenge for theoretical hydrology', *Hydrological Processes*, 17: 3163-70.
- Smith, Laurence C., Donald L. Turcotte, and Bryan L. Isacks. 1998. 'Stream flow characterization and feature detection using a discrete wavelet transform', *Hydrological Processes*, 12: 23349.
- Stella, John C, John J Battles, Bruce K Orr, and Joe R McBride. 2006. 'Synchrony of seed dispersal, hydrology and local climate in a semi-arid river reach in California', *Ecosystems*, 9: 1200-14.
- Stewart-Koster, Ben, Julian D. Olden, and Keith B. Gido. 2014. 'Quantifying flow–ecology relationships with functional linear models', *Hydrological Sciences Journal*, 59: 629-44.
- Wagener, T., Sivapalan, M. 2007. 'Catchment Classification and Hydrologic Similarity', *Geography Compass*.
- Webb, J. Angus, Kimberly A. Miller, Elise L. King, Siobhan C. de Little, Michael J. Stewardson, Julie K. H. Zimmerman, and N. LeRoy Poff. 2013. 'Squeezing the most out of existing literature: a systematic re-analysis of published evidence on ecological responses to altered flows', *Freshwater Biology*, 58: 2439-51.
- Wigington, P., Leibowitz, S. 2012. 'Oregon Hydrologic Landscapes: A Classification Framework'.
- Wilcoxon, Frank. 1945. 'Individual comparisons by ranking methods', *Biometrics bulletin*, 1: 8083.
- Wolock, David M, Thomas C Winter, and Gerard McMahon. 2004. 'Delineation and evaluation of hydrologic-landscape regions in the United States using geographic information system tools and multivariate statistical analyses', *Environmental Management*, 34: S71-S88.
- Yang, Han-Chung, Jian-Ping Suen, and Shih-Kai Chou. 2016. 'Estimating the Ungauged Natural Flow Regimes for Environmental Flow Management', *Water Resources Management*, 30: 4571-84.
- Yarnell, Sarah M., Geoffrey E. Petts, John C. Schmidt, Alison A. Whipple, Erin E. Beller, Clifford N. Dahm, Peter Goodwin, and Joshua H. Viers. 2015. 'Functional Flows in Modified Riverscapes: Hydrographs, Habitats and Opportunities', *BioScience*, 65: 963-72.
- Yarnell, Sarah M., Joshua H. Viers, and Jeffrey F. Mount. 2010. 'Ecology and Management of the Spring Snowmelt Recession', *BioScience*, 60: 114-27.

Beyond Metrics

Supplemental Materials

Metastudy of existing hydrologic classifications

Metastudies synthesize multiple datasets and viewpoints on a scientific topic. They are common in health and social sciences (e.g., (Hunt 2006), but uncommon in geosciences, including hydrology. One reason is that it is rare for multiple geoscience studies to use the same data or do highly similar analyses of the same setting. Even efforts to reconcile scientific concepts across diverse studies within a broad theme are relatively rare, though much needed (Blöschl 2006). For the purposes of this study, the ability to reconcile overlapping, spatially explicit hydrological models of archetypal baselines for a region to develop a single parsimonious model representing the best features of each effort would be a major contribution to the scientific literature. To the best of the authors’ knowledge, no such metastudy has been completed to date for any region. Scientific and management applications for in-stream flow allocations are expected to benefit considerably from a unified consensus stream classification system capable of supporting a broad range of programs aimed at protecting water quality, ecological health, and adequate water supply for consumptive and non-consumptive uses.

The first phase of this study involved reconciling two existing hydrologic classifications of California in order to provide a single consensus classification system. These two efforts (Pyne et al. 2017; Lane et al. 2017) addressed different water management objectives, yet arrived at similar, though not totally identical, results. The similarities attest to the resilience of the underlying data in providing a clear signal that drove convergent outcomes.

The two classifications exhibited distinct data inputs, classification and prediction methods, and resulting stream classes (Table 1). In terms of hydrologic input data, Lane et al. (2017) evaluated significantly fewer reference gauge stations, but included a broader range of hydrologic patterns by incorporating naturalized as well as unimpaired stream gages (Figure 1). Streamflows were naturalized by manually calculating the influence of impairments and removing this influence to reflect expected conditions in the absence of impairment. The reference gauges used by Lane et al. (2017) were also unevenly distributed across the state, with the vast majority of gauges on tributaries to the Sacramento and San Joaquin watersheds and minimal gauges in southern California (Figure 1). In contrast, the gauges used by Pyne et al. (2017) had a more even distribution across the state, but were restricted to gauges draining “unaltered” catchments. Thus, each approach was sensible for its original purpose and taken together there is an opportunity for reconciliation among them.

Table 1. Comparison table of two hydrologic classification studies for California, including data inputs, methods, and results.

	<i>Lane et al. (2017)</i>	<i>Pyne et al. (2017)</i>
<i># of reference gauges</i>	91, unimpaired and naturalized	138, unimpaired
<i>Classification, Method</i>	Hydrologic metrics, K-means non-hierarchical and Ward’s hierarchical clustering	Physical characteristics, Unsupervised Bayesian modeling and Ward’s hierarchical clustering

<i>Prediction/validation, Method</i>	Physical characteristics, Classification and Regression Trees	Hydrologic metrics, Random Forests
<i>Number of classes</i>	8	7
<i>Misclassification rate</i>	14%	7 – 21%
<i>Dominant hydrologic metrics</i>	Seasonality, inter-annual variability, annual magnitude	High flow magnitude, flow recession, persistence
<i>Dominant physical characteristics</i>	Elevation, contributing area, winter precipitation	Winter precipitation, geology, elevation

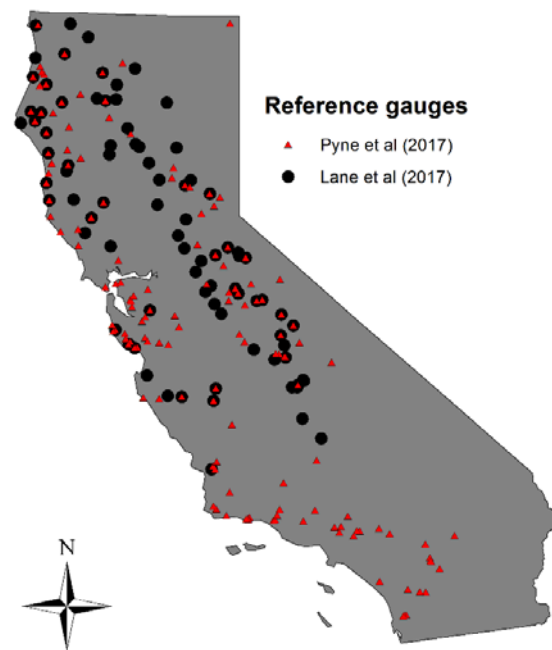


Figure 1. Distribution of reference streamflow gauge stations used by Lane et al. (2017) and Pyne et al. (2017).

While different statistical tools were employed, these two studies applied essentially inverted methods. They are inverted in the sense that one approach started with hydrological data and then analyzed the catchment attributes of clustered time series, while the other started with catchment data and then analyzed the hydrological attributes of clustered physical entities. Specifically, Lane et al. (2017) used flow metrics extracted from hydrological time series as inputs to a cluster analysis and then predicted the spatial distribution of the resulting stream classes based on physical catchment characteristics. Alternatively, Pyne et al. (2017) used physical catchment characteristics as inputs to a landscape classification model and then evaluated hydrologic groupings resulting from the classification.

Dominant physical catchment characteristics identified by each study were quite similar, with both classifications distinguishing elevation and winter precipitation as key controls on hydrologic response. However, the order of significance and other key physical controls, such as geology and drainage area, differed between the two studies. Conceptually, all of the key

physical controls across both studies are widely recognized as important in differentiating hydrological regimes where there is a gradient for each variable in a dataset.

In contrast, the identified dominant hydrologic metrics in each study differed dramatically, with inter- and intra-annual flow variability metrics dominating in Lane et al. (2017) while high flow magnitude and flow duration dominated in Pyne et al. (2017). As the studies did not use all the same flow metrics, this difference in dominant indices, which has major management implications, is due in part to a neglect of some significant flow metrics by each study. We hypothesize that the additional use of hydrologic archetypes will inform the development of a hydrologic conceptual model or the selection of flow metrics such that the resulting hydrologic baselines will provide a more complete understanding of ecologically relevant hydrologic variability.

Final reconciled hydrologic classifications

Figure 2 details the final stream class resulting from dominant combinations of Lane et al (2017) and Pyne et al (2017). For instance, when SM streams overlap with Class 6 streams, they will be classified as SM as they fall under outcome c in the reconciliation workflow (Step 1, Figure 2) because they are in agreement. Alternatively, when LSR streams overlap with Class 2 streams, they will be classified as HLP. This subregion of disagreement led to outcome b in the workflow and resulted in splitting LSR into a new class when it overlapped with Class 2 due to its physical distinction from LSR. The asterisks indicate that the resulting classification is only true in a particular region. For instance, RGW and Class 3 will only lead to FER in the Mohave region; otherwise, RGW stream class will be retained.

Lane et al. (2017)	Pyne et al. (2017)	Final Classification
SM	Class 4 or 6	SM
HSR	Class 7	HSR
LSR	Class 1 or 4	LSR
	Class 5	FER
	Class 3	HLP*
	Class 5	RSG
RSG	Class 3	FER*
	Class 4	WS
WS	Class 4	WS
GW	Class 6 or 7	GW
PGR	Class 1 or 2	PGR
FER	Class 5	FER

Figure 2. Reconciliation process detailing the final classification resulting from the reconciliation of major subregions of agreement and disagreement between Lane et al. (2017) and Pyne et al. (2017). Asterisks indicate that the final stream classes only result in particular subregions. For instance, RGW and Class 3 will only lead to FER in the Mohave region. Otherwise, RGW (from Lane et al. 2017) is retained.

The distribution of stream lengths by stream class (Figure 3) indicates that the majority of California streams by length naturally exhibit low-volume snowmelt and rain (LSR) and rain and seasonal groundwater (RGW) hydrologic regimes. Alternatively, the least common hydrologic regimes are groundwater-dominated and high-volume snowmelt and rain.

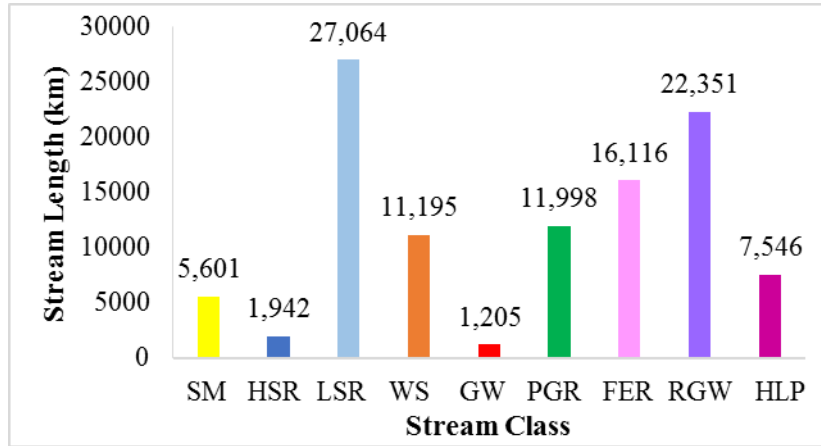











Figure 3. Stream length distribution of final stream classes

Corrected

Table 2. Summary table of reconciled hydrologic classification for the State of California

Class	Name	Hydrologic Characteristics	Physical and Climatic Catchment Controls
	Snowmelt	<ul style="list-style-type: none"> • Large spring snowmelt pulse (~May 24) • Very high seasonality index • Extreme low flows (<10th percentile) Sep-Feb 	<ul style="list-style-type: none"> • High elevation catchments (>2,293 m), major snow influence
	Low-volume snowmelt and rain	<ul style="list-style-type: none"> • Transition between SM and HSR • Bimodal snow - rain hydrograph driven by spring snowmelt pulse and winter rain 	<ul style="list-style-type: none"> • Mid-elevation catchments with limited area (<2,144 km²) [low winter temperatures (Jan temp <-5°C), high stream density (>0.65 km/km²)]
	High-volume snowmelt and rain	<ul style="list-style-type: none"> • Spring snowmelt pulse (~May 4) • High seasonality but larger winter storm contributions • Retain high baseflow throughout summer • Bimodal snow - rain hydrograph 	<ul style="list-style-type: none"> • Mid-elevation catchments (1,126 - 2,293 m), large contributing area (>2,144 km²) <i>not</i> underlain by volcanic geology [high stream density (>0.65 km/km²), mild winter temperatures (Jan temp >-5°C)] OR • Low elevation (<1,125 m) with very large contributing area (>15,420 km²) and high clay content soils (>17% clay)
	Winter storms	<ul style="list-style-type: none"> • Predictable large fall and winter storms • Earliest peak flows (in January) 	<ul style="list-style-type: none"> • Low elevation catchments with substantial winter precipitation OR • Low elevation, mid-slope (31 - 24%) catchments with low winter precipitation but high riparian soils clay content (>23%) AND • Underlain by unconsolidated aquifers covered by thick alluvium
	Groundwater	<ul style="list-style-type: none"> • Highest mean annual flows and highest minimum flows • Low seasonality and high predictability 	<ul style="list-style-type: none"> • Mid-elevation catchments with large area (>2,144 km²) underlain by volcanic (basaltic and andesitic) geology [low stream density (<0.65)] OR • Low elevation, limited winter precipitation, very large contributing area (>15,420 km²) with low riparian soils clay content (<17%) AND • Underlain by igneous and metamorphic rock aquifers
	Perennial groundwater and rain	<ul style="list-style-type: none"> • Low seasonality and mean annual streamflow • Transition between WS and GW, with winter rain contributions but generally stable flows 	<ul style="list-style-type: none"> • Low elevation catchments with low clay content riparian soils (<23%) [low stream density (<1.1 km/km²)] AND • Underlain by sedimentary rock materials in Central Coast region
	Rain and seasonal groundwater	<ul style="list-style-type: none"> • Bimodal hydrograph driven by winter rain pulse and percolating winter rain appearing as baseflow pulse later in year; can be ephemeral 	<ul style="list-style-type: none"> • Low elevation with limited winter precipitation and low slopes (<24%) AND • Coastal catchments with small aquifers driving short residence times
	Ephemeral, flashy rain	<ul style="list-style-type: none"> • Lowest mean annual flows, often ephemeral • Highest CV, lowest predictability • Longest extreme low flow duration • Highest flows in winter 	<ul style="list-style-type: none"> • Low elevation catchments with high clay content soils (>23%) and high slopes (>31%) [high stream density (>1.15 km/km²)]
	High elevation, low precipitation	<ul style="list-style-type: none"> • Low mean annual flows • Highest flows in winter 	<ul style="list-style-type: none"> • High elevation but low slope • Low precipitation and limited snow influence

Stream class RBI comparisons (Table 3) in particular distinguish broad-scale differences in flow flashiness, ranging from an average of 0.06 in HLP to 0.55 in FER. Not surprisingly, the ephemeral FER stream class was the flashiest, followed closely by winter rain dependent PGR, RGW, and WS classes. LSR and SM are decreasingly flashy as they are increasingly snowmelt dominated. LSR still exhibits a relatively high RBI due to the evident rain influence, while SM is extremely stable (RBI=0.12). The HLP class, characterized by stable groundwater-driven flows and limited snowmelt influence, is the least flashy (RBI=0.06). However, these results display a large spread, making it difficult to distinguish between all but the most distinctive stream classes. Further, no seasonal comparisons of RBI values are possible because metrics were calculated over the entire population of streamflow data rather than retaining daily timing attributes.

Table 3. Stream class RBI comparisons quantify difference in stream class flashiness.

Class	RBI			
	avg	std	min	max
HLP	0.06	0.05	0.03	0.10
SM	0.12	0.04	0.04	0.20
LSR	0.27	0.14	0.03	0.67
WS	0.40	0.13	0.13	0.85
RGW	0.48	0.20	0.14	0.80
PGR	0.48	0.17	0.04	0.85
FER	0.55	0.18	0.30	0.90

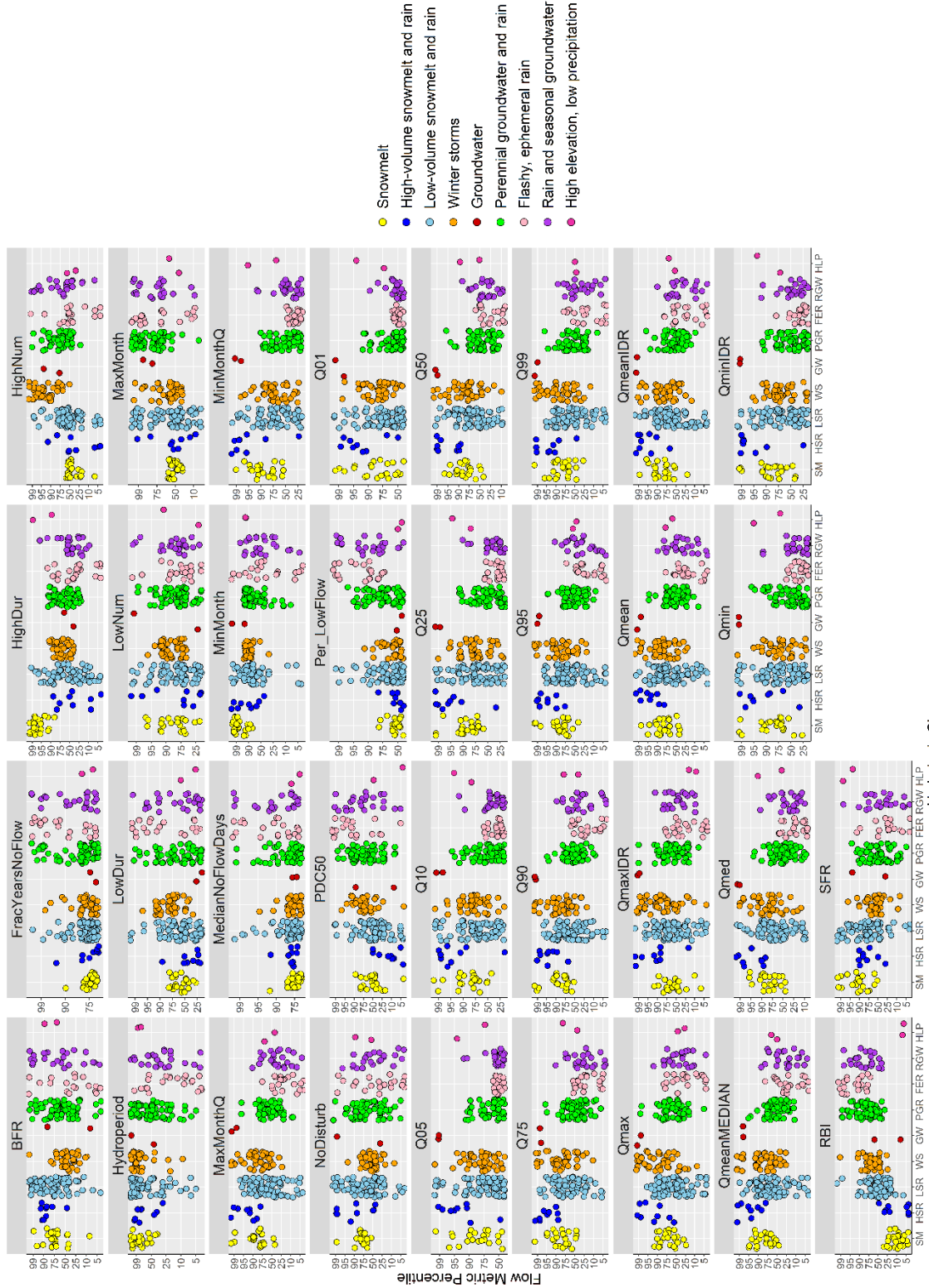


Figure 4. 34 common flow metrics (from Konrad et al. 2008) calculated for each reference gage in each stream class.

- Blöschl, Günter. 2006. 'Hydrologic synthesis: Across processes, places, and scales', *Water Resources Research*, 42.
- Hunt, Terry L. 2006. 'Rethinking the fall of Easter Island'.
- Konrad, CP, AMD Brasher, and JT May. 2008. 'Assessing streamflow characteristics as limiting factors on benthic invertebrate assemblages in streams across the western United States', *Freshwater Biology*, 53: 1983-98.
- Lane, Belize A., Helen E. Dahlke, Gregory B. Pasternack, and Samuel Sandoval-Solis. 2017. 'Revealing the Diversity of Natural Hydrologic Regimes in California with Relevance for Environmental Flows Applications', *JAWRA Journal of the American Water Resources Association*: n/a-n/a.
- Pyne, Matthew, D. M. Carlisle, C. P. Konrad, and Eric Stein. 2017. 'Classification of California Streams', *Ecohydrology*.

Corrected, final