Appendix H: An assessment of streamflow alteration status and biological conditions in California streams

Authors: Ryan Peek[1], Sarah Yarnell[1], Rob Lusardi[1,2], Eric Stein[3], Katie Irving[3], Raphael Mazor[3]
[1] Center for Watershed Sciences, University of California, Davis
[2] CalTrout
[3] Southern California Coastal Water Research Project

Introduction

A critical component of developing ecological flow criteria is quantifying the extent and degree of streamflow alteration in relation to biological condition data. The California Stream Condition Index (CSCI) quantifies biological condition at a given sampling site or stream segment by translating benthic macroinvertebrate and watershed-scale environmental data into an overall measure of stream health (Mazor et al. 2016). CSCI provides a consistent statewide standard for interpreting bioassessment data and thus a means of quantitatively and objectively comparing stream condition throughout the state.

The objective of this analysis was to compare streamflow alteration status, quantified as described in Appendix F, with biological condition, quantified by CSCI, across California. Specifically, we assessed functional flow metrics that best correlated with stream health condition derived from paired BMI (benthic macroinvertebrate) stream surveys and USGS sites at two scales (regionally and statewide). The results of this analysis can be used in Section 2 of the CEFF guidance document to prioritize functional flow metrics (and respective associated response relationships) when developing ecological flow criteria.

Commented [SY1]: Simplified to reflect our most recent discussions of how we're going to update the main CEFF guidance doc.

Background

Benthic macroinvertebrates have been successfully used as indicators of stream health in a wide range of studies (Lawrence et al. 2010; Lunde et al. 2013; Mazor et al. 2016). Notably, they have been extensively used to quantify biological impairment associated with shifts in the environment. For instance, hydrologic alteration or impairment has been shown to strongly influence aquatic communities (Poff et al. 2007; Rehn 2009), and benthic macroinvertebrates have more recently been used to link metrics of hydrologic variability to biological response (Poff et al. 2010; Steel et al. 2017).

The Surface Water Ambient Monitoring Program (SWAMP) is tasked with assessing surface water quality throughout California. The program coordinates water quality monitoring across the state and collects data to support water resource management by the Water Boards. For example, the data collected by SWAMP's probabilistic Perennial Stream Assessment survey is used to characterize in-stream biological condition and make estimates about the extent of healthy streams in different regions of the state. These data include several biological indicators, including BMI, benthic algae, and measures of physical habitat integrity. These data can be used to calculate the CSCI and other quantitative measures of stream condition. Leveraging this statewide dataset in conjunction with new tools for quantifying hydrologic variability at the stream segment scale across California, as presented in CEFF, provides a unique opportunity to assess biological response to hydrologic alteration in California.

Methods

To begin exploring relationships between streamflow conditions and stream health, we statistically compared observed CSCI scores with modeled estimates of flow alteration status for functional flow metrics calculated at nearby USGS gage stations. We spatially and temporally paired existing stream bioassessment survey data with proximal USGS gage data. We calculated flow alteration status, as described in appendix F, for 24 functional flow metrics defined by Yarnell et al. (2020) at each USGS site paired with a bioassessment site. We developed statistical models to identify which of the 24 functional flow metrics were most closely associated with CSCI scores. Additional details on each step in the analysis are provided below.

Pairing of CSCI data with USGS gage data

We identified all (n=2,935) bioassessment sites in the SWAMP dataset with available CSCI scores from data sampled between 1994-2018 during late spring and summer months (May to September, when sampling typically occurs). To pair these BMI sites with USGS gages, we filtered the list of BMI locations to include only sites that occurred in the same HUC12 sub-watershed as USGS sites with at least 10 years of daily flow data (Figure 1). We then manually filtered the BMI sites to include only those that occurred on the same National Hydrography Dataset (NHD) mainstem streamline as the USGS gage and were within 10 km of the gage using the nhdplusTools, dplyr, and sf packages in R version 3.5.3 (R Core Team 2019; Wickham et al. 2018; Blodgett 2018; Pebesma 2018) . Using this list of BMI-USGS site pairs, we then filtered site pairs to only include flow data post-1994 to ensure temporal overlap with the BMI sampling events. Data from these final site pairs were used in all subsequent analyses.

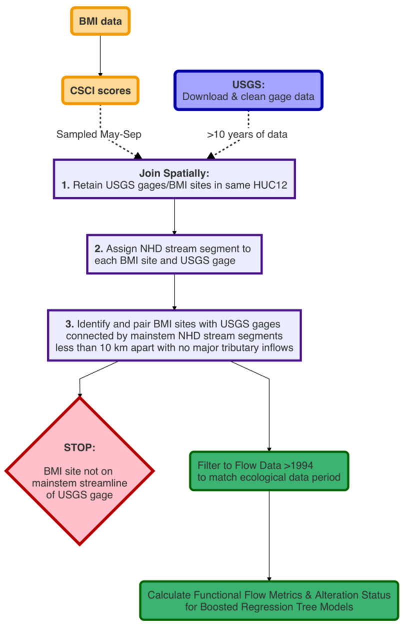

Figure 1. Flow diagram of steps used to pair BMI and USGS datasets. Note, currently there is no public statewide dataset of CSCI scores.

Determination of Flow Alteration Status

Once the selected CSCI data was paired with relevant USGS data, we calculated functional flow metrics over the period of record for each USGS gage using the functional flows calculator as described in Appendix D. Following the steps delineated in Appendix F to determine the flow alteration status at each gage, we then compared the observed functional flow metrics to the predicted functional flow metrics associated with the stream segment at the USGS gage (see Appendix E for additional details on how predicted reference-based functional flow metrics were modeled). Any observed functional flow metric that was determined to be “Likely Unaltered” (see Appendix F) was assigned an alteration status score of 1, metrics that were determined to be “Likely Altered” were assigned a score of -1, and metrics that were determined to be “Indeterminate” were assigned a score of 0. In some cases, the functional flow metric value for a single water year at a gage could not be calculated, resulting in an ‘NA’ value for that year. If more than 70% of the annual values for a metric across the period of record at a gage were NA, then the flow alteration for that metric at that gage was not included in the dataset. One metric, spring recession duration, had to be dropped from analysis because it was missing more than 70% of the annual values for all gages. Thus, for every site pair, data included a single CSCI score and up to 23 flow alteration status scores, one for each of the remaining functional flow metrics.

Statistical Analysis using Boosted Regression Trees

In order to determine which functional flow metrics were most associated with a gradient of streamflow alteration, we modeled flow alteration status estimates for each functional flow metric against CSCI scores using boosted regression tree analysis, following methods from Steel et al. (2017). Boosted regression trees, a method from the decision tree family of statistics, are well suited for large and complex ecological datasets; they do not assume normality nor linear relationships between predictor and response variables, they ignore non-informative predictor variables, and they can accept predictors that are numeric, categorical, or binary (Brown et al. 2012; Elith, Leathwick, and Hastie 2008). Boosted regression trees are also unaffected by outliers and effectively handle both missing data and collinearity between predictors (De'ath 2007; Dormann et al. 2013). Importantly, such methods are becoming more common in ecological analyses and have been shown to outperform many traditional statistical methods such as linear regression, generalized linear models, and generalized additive models (Guisan et al. 2007). The boosted regression tree models were run with grid iteration and tuning across parameters (shrinkage [0.001–0.009], interaction depth [3–5], number of minimum observations in a node [3–10], and bag fraction [0.75–0.8]) in model validation. To assess the relative influence of each functional flow metric in the model, we used the mean-square error method (Ridgeway 2015).

The top two most influential functional flow metrics were further examined by plotting the raw metric values against CSCI scores. In order to see if these relationships were consistent in different regions, we completed this analysis at both the statewide and regional scale. Geographic regions encompassed the HUC4 watershed scale (see Appendix G) and included the Central Valley region (including the west slope of the Sierra Nevada), North Coast region, Great Basin region, and South Coast region (Figure 2).

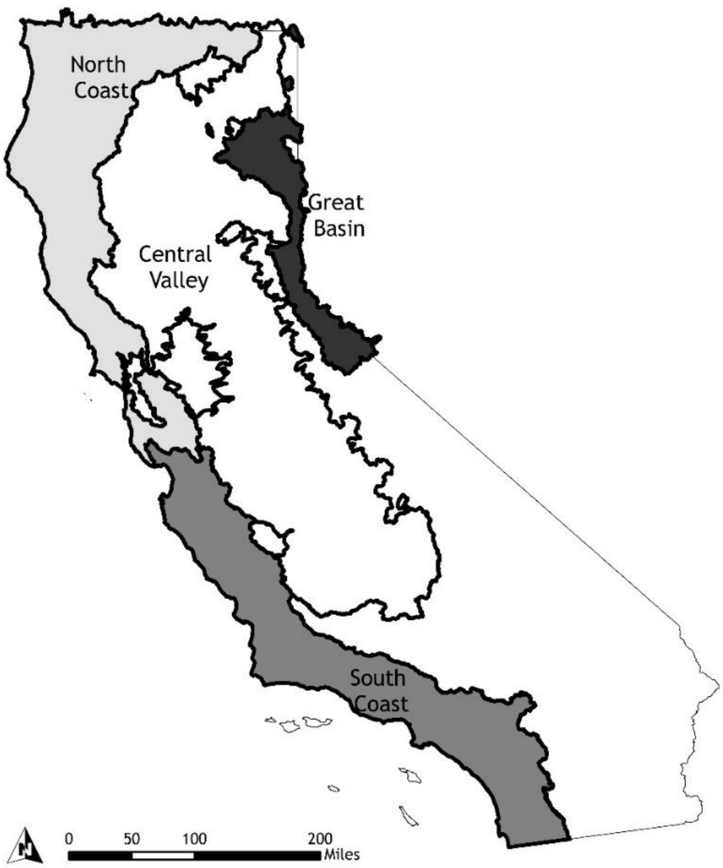

Figure 2. Geographic regions in California, as defined in Appendix G.

Results

Pairing of CSCI data with USGS gage data

We mapped a total of 2,935 unique BMI sampling locations and 799 USGS gage sites (Figure 3) across California. After pairing sites, we identified 270 BMI sampling sites associated with 160 USGS gages across the state. In some cases, CSCI scores were from data collected over different years at the same site or more than one sampling site was associated with a nearby USGS gage. For BMI sampling events and resulting CSCI scores that occurred in the same water year at the same location, we calculated the median value of these replicate CSCI scores to use in the statistical modeling.

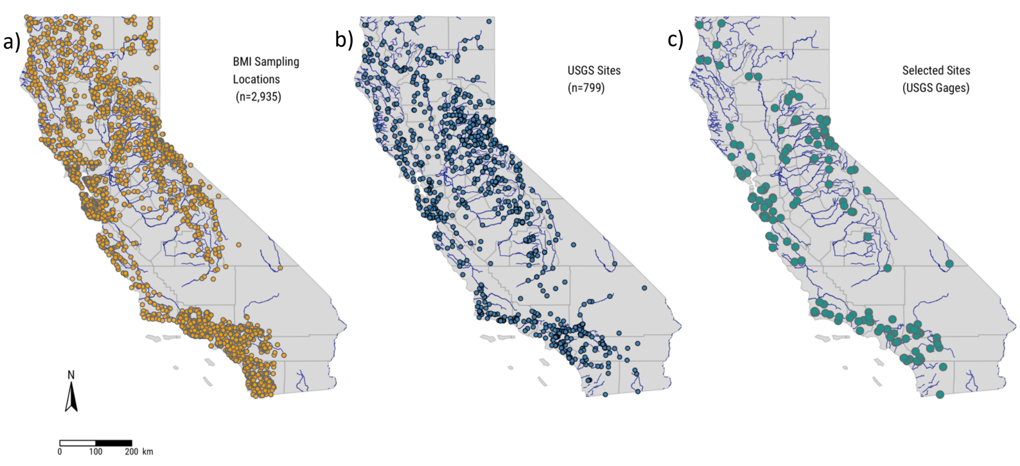

Figure 3. Map of all BMI sampling sites (a), all USGS gages sites with >10 years of flow data (b), and the selected gages (n=160) for the final site pairs (c).

Statistical Analysis for Statewide Site Pair Dataset

Boosted regression tree models were run using all site paired data across the state to determine the functional flow metrics most associated with streamflow alteration in relation to CSCI scores. Mean-square error was used to calculate variable influence of functional flow metrics. The final model explained 21% of the deviance and had a cross-validation correlation of 0.603 (standard error, 0.133). Of the 23 functional flow metrics included in the model

Commented [ES2]: Redundant with above text

Commented [SY3R2]: I suggest leaving to ensure clarity.

(spring recession duration was not included due to inadequate data), six had relative importance values greater than 5% (Figure 4, Table 1).

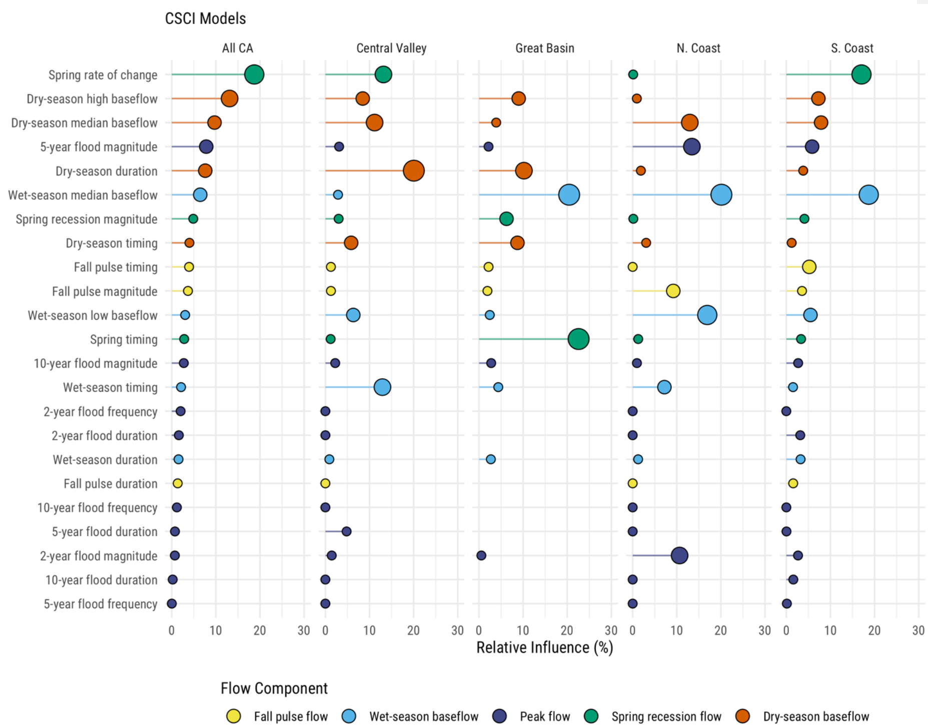

Figure 4. Relative importance of functional flow metrics in boosted regression tree models assessing flow alteration status relative to CSCI score for paired sites statewide and regionally. Relative influence values were calculated using a mean-square error (MSE) approach, which determines those variables with the largest average reduction in MSE. Functional Flow Metrics are described in Table 1.

Table 1. Mean Relative Influence values based on mean square error for functional flow metrics included in the statewide model assessing flow alteration status in relation to CSCI score. Metrics and relative influence values in bold were most influential (> 5%).

Flow Component	Flow Metric Name (unit)	Flow Metric Description	Relative Influence (%)
Dry-season baseflow	Dry-season timing (water year day)	Dry-season baseflow start timing (water year day of dry season)	4
Dry-season baseflow	Dry-season high baseflow (cfs)	90th percentile of daily flow within dry season	**13.1**
Dry-season baseflow	Dry-season median baseflow (cfs)	50th percentile of daily flow within dry season	**9.7**
Dry-season baseflow	Dry-season duration (days)	Dry-season baseflow duration (# of days from start of dry season to start of wet season)	**7.6**
Fall pulse flow	Fall pulse timing (water year day)	Water year day of fall pulse event peak	3.9
Fall pulse flow	Fall pulse magnitude (cfs)	Peak magnitude of fall pulse event (maximum daily peak flow during event)	3.7
Fall pulse flow	Fall pulse duration (days)	Duration of fall pulse event	1.4
Peak flow	5-year flood magnitude (cfs)	5-year recurrence interval flood flow	**7.8**
Peak flow	2-year flood magnitude (cfs)	2-year recurrence interval flood flow	0.7
Peak flow	10-year flood magnitude (cfs)	10-year recurrence interval flood flow	2.7
Peak flow	5-year flood duration (days)	Seasonal duration of 5-year recurrence interval flood flow (cumulative number of days 5-year flood flow is exceeded)	0.7
Peak flow	2-year flood duration (days)	Seasonal duration of 2-year recurrence interval peak flow (cumulative number of days 2-year flood is exceeded)	1.6
Peak flow	10-year flood duration (days)	Seasonal duration of 10-year recurrence interval peak flow (cumulative number of days 10-year flood flow is exceeded)	0.2
Peak flow	5-year flood frequency (occurrences)	Frequency of 5-year recurrence interval flood flow within a season	0
Peak flow	2-year flood frequency (occurrences)	Frequency of 2-year recurrence interval flood flow within a season	2
Peak flow	10-year flood frequency (occurrences)	Frequency of 10-year recurrence interval flood flow within a season	1.2
Spring recession flow	Spring timing (water year day)	Start date of spring in water year days	2.8
Spring recession flow	Spring rate of change (percent)	Spring flow recession rate (median daily rate of change over	**18.7**

Flow Component	Flow Metric Name (unit)	Flow Metric Description	Relative Influence (%)
		decreasing periods during the recession)	
Spring recession flow	Spring recession magnitude (cfs)	Spring recession magnitude (daily flow on start date of spring-flow period, 4 days after last wet-season peak)	4.9
Wet-season baseflow	Wet-season timing (water year day)	Start date of wet-season in water year days	2.1
Wet-season baseflow	Wet-season median baseflow (cfs)	Magnitude of wet-season baseflows (50th percentile of daily flows within that season, including peak flow events)	**6.5**
Wet-season baseflow	Wet-season low baseflow (cfs)	Magnitude of wet-season baseflows (10th percentile of daily flows within that season, including peak flow events)	3
Wet-season baseflow	Wet-season duration (days)	Wet-season baseflow duration (# of days from start of wet-season to start of spring season)	1.5

The two most influential functional flow metrics in the statewide model included the spring recession rate of change (18.7 % relative influence) and the 90th percentile of the dry season baseflow magnitude (11.1% relative influence) (Table 1). These two metrics were further examined by plotting the values for each functional flow metric (instead of alteration status) against paired CSCI score (Figures 5-6). Both plots show the highly variable nature of the large dataset, which is to be expected given the inherent wide diversity of sites across California. However, trends in the data indicate potential underlying relationships that should be explored further. Using a general additive model (plotted as a trend line in each figure), the data indicate that as the spring recession rate of change increases, the CSCI score decreases (Figure 6). Conversely, for the 90th percentile dry season baseflow magnitude, CSCI score decreases at high discharges and is highest at moderate discharges (Figure 8).

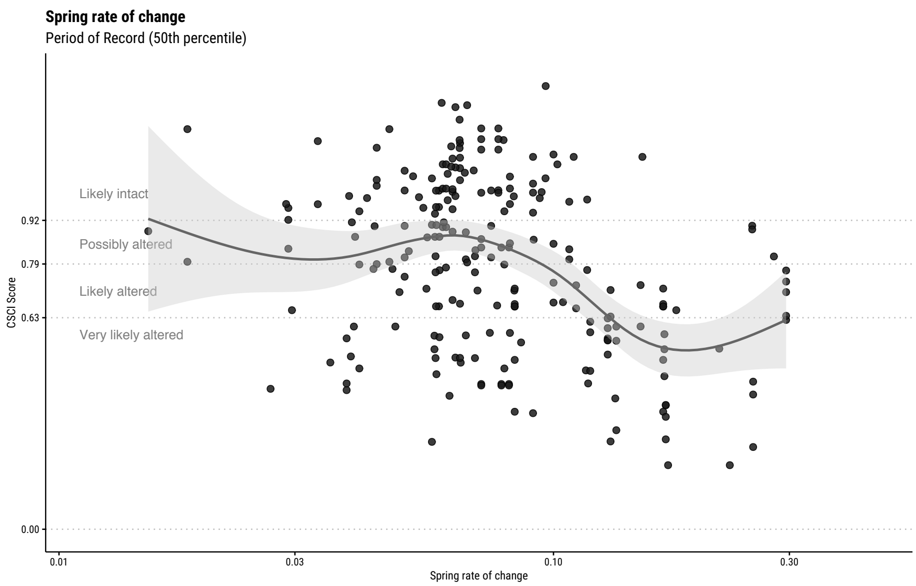

Figure 5. Median spring recession rate of change values from the period of record plotted against CSCI scores for all site pairs in the statewide dataset. The trend line shows a general additive model fit to the data indicating that as spring recession rate of change increases, CSCI score decreases. The y-axis denotes CSCI thresholds of biological condition based on Mazor et al. (2016). CSCI scores ≥ 0.92 indicate streams are likely intact, ≥ 0.79 indicate possibly altered, ≥ 0.63 indicate likely altered, and scores < 0.63 indicate stream condition is very likely to be altered. Note that the x-axis is plotted on a logarithmic scale, while the y-axis is plotted on an interval scale.

Commented [SY4]: So are we leaving the values in behind the trendline or leaving it as shown with just the trendline? If leaving as shown, then need to edit the first sentence accordingly. If leave as shown, then also could adjust the plot to have less white space above and below the trendline. I think I like seeing the data behind the trendline, but can go either way.

Commented [RAP5R4]: Important to keep the white space above and below because it shows range of data (and need to show zero). We need to keep the points in as well.

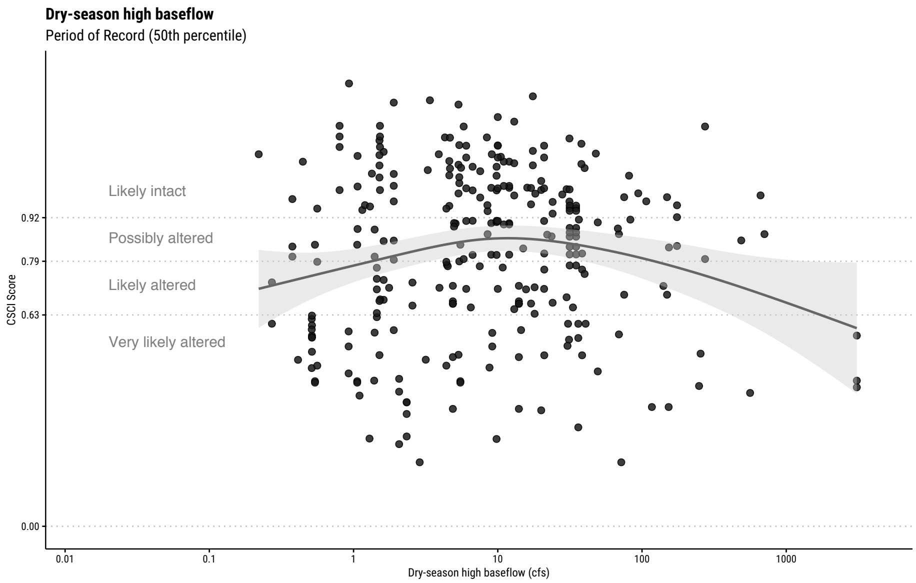

Figure 6. Median dry-season high (90th percentile) baseflow magnitude values plotted against CSCI scores for all site pairs in the statewide dataset. The trend line shows a general additive model fit to the data indicating that at moderate discharges, CSCI scores are highest, and at high discharges, CSCI score decreases. The y-axis denotes CSCI thresholds of biological condition based on Mazor et al. (2016). CSCI scores ≥ 0.92 indicate streams are likely intact, ≥ 0.79 indicate possibly altered, ≥ 0.63 indicate likely altered, and scores < 0.63 indicate stream condition is very likely to be altered. Note that the x-axis is plotted on a logarithmic scale, while the y-axis is plotted on an interval scale.

Statistical Analysis for Regional Site Pair Datasets

To better understand regional patterns, we analyzed CSCI scores and estimated flow alteration status using boosted regression tree models for each of the geographic regions defined in Appendix G (Figure 2). Results suggest that wet season and dry season flow components were important in all regional models, while the fall pulse flow and peak magnitude flow components were more influential in the North Coast and South Coast regional models (Table 2). For the Central Valley regional model, all dry-season functional flow metrics were influential, along with the spring recession rate of change, wet season timing, and the 10th percentile of wet season baseflow magnitude (Figure 4). For the Great Basin regional model, dry season functional flow metrics were influential, as well as spring recession timing and magnitude and wet season median magnitude (Figure 4). For the North Coast and South Coast regional models, fall pulse timing and magnitude were influential as well as the 2-year and 5-year peak flow magnitudes (Figure 4). Similar to the other regional models, dry season and wet season baseflow magnitudes were also influential in the North and South Coast regional models.

Table 2. Mean Relative Influence values for functional flow metrics included in each of the four regional models that assessed flow alteration status in relation to CSCI score. Metrics and relative influence values in bold were most influential (> 5%). Relative influences denoted as 'NA' indicate that functional flow metric was not included in the regional model due to lack of data.

Flow Component	Flow Metric Name (unit)	Central Valley (RI %)	Great Basin (RI %)	North Coast (RI %)	South Coast (RI %)
Dry-season baseflow	Dry-season timing (water year day)	**5.8**	**8.7**	3.1	1.2
Dry-season baseflow	Dry-season high baseflow (cfs)	**8.5**	**9**	0.9	**7.3**
Dry-season baseflow	Dry-season median baseflow (cfs)	**11.2**	3.9	**12.9**	**7.9**
Dry-season baseflow	Dry-season duration (days)	**20**	**10.2**	1.9	3.8
Fall pulse flow	Fall pulse timing (water year day)	1.3	2.2	0	**5.2**
Fall pulse flow	Fall pulse magnitude (cfs)	1.3	1.9	**9.2**	3.6
Fall pulse flow	Fall pulse duration (days)	0	NA	0	1.6
Peak flow	5-year flood magnitude (cfs)	3.1	2.2	**13.4**	**5.8**
Peak flow	2-year flood magnitude (cfs)	1.4	0.5	**10.6**	2.7
Peak flow	10-year flood magnitude (cfs)	2.2	2.8	1	2.7
Peak flow	5-year flood duration (days)	4.8	NA	0	0
Peak flow	2-year flood duration (days)	0	NA	0	3.2
Peak flow	10-year flood duration (days)	0	NA	0	1.6
Peak flow	5-year flood frequency (occurrences)	0	NA	0	0.1
Peak flow	2-year flood frequency (occurrences)	0	NA	0	0
Peak flow	10-year flood frequency (occurrences)	0	NA	0	0
Spring recession flow	Spring timing (water year day)	1.2	**22.6**	1.3	3.3
Spring recession flow	Spring rate of change (percent)	**13.2**	NA	0.1	**17**
Spring recession flow	Spring recession magnitude (cfs)	3	**6.2**	0.2	4.1

Flow Component	Flow Metric Name (unit)	Central Valley (RI %)	Great Basin (RI %)	North Coast (RI %)	South Coast (RI %)
Wet-season baseflow	Wet-season timing (water year day)	**12.9**	4.4	**7.2**	1.5
Wet-season baseflow	Wet-season median baseflow (cfs)	2.9	**20.4**	**20.1**	**18.7**
Wet-season baseflow	Wet-season low baseflow (cfs)	**6.3**	2.4	**16.9**	**5.5**
Wet-season baseflow	Wet-season duration (days)	0.9	2.7	1.2	3.2

Closer examination of the values of several key functional flow metrics relative to CSCI score within each region showed similar trends to statewide model results. Trend lines fit to the data indicated that as spring recession rate of change increased in all regional models (except Great Basin), CSCI score decreased (Figure 7). For the 90th percentile of dry season baseflow magnitude values, trend lines indicated that CSCI scores were highest at moderate discharges, while CSCI decreased at high discharge (Figure 8). However, for the median wet season baseflow magnitude, trend lines varied across regional models, such that moderate discharges correlated with higher CSCI scores in the Central Valley and North Coast regions, but lower wet season median discharges correlated with higher CSCI scores in the Great Basin region (Figure 9). There was no trend with wet season median discharge in the South Coast region (Figure 9).

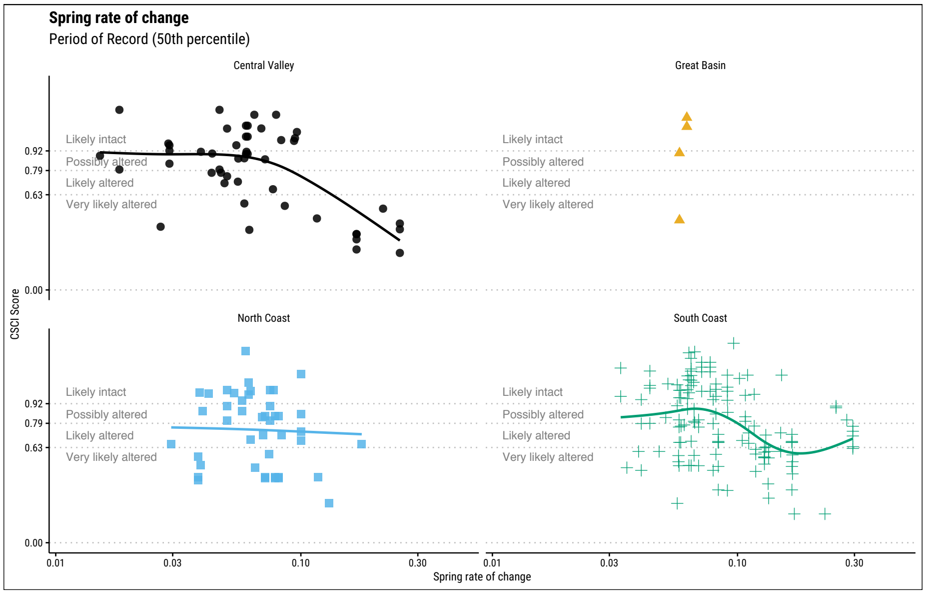

Figure 7. Trend in median spring recession rate of change values from the period of record relative to CSCI scores for all site pairs in each regional dataset. The trend lines show a general additive model fit to the data indicating that as spring recession rate of change increases in Central Valley and South Coast, CSCI score decreases. Note that the x-axis is plotted on a logarithmic scale, while the y-axis is plotted on an interval scale.

Commented [SY6]: I think for these plots, don't show the values behind the trend lines and just describe the trendline (not the values behind the trendline).

Commented [RAP7R6]: I think we need to keep the values in…particularly because if you don't we'll just have a blank panel in one, and it helps show how much data exists for each region and thus the trendline.

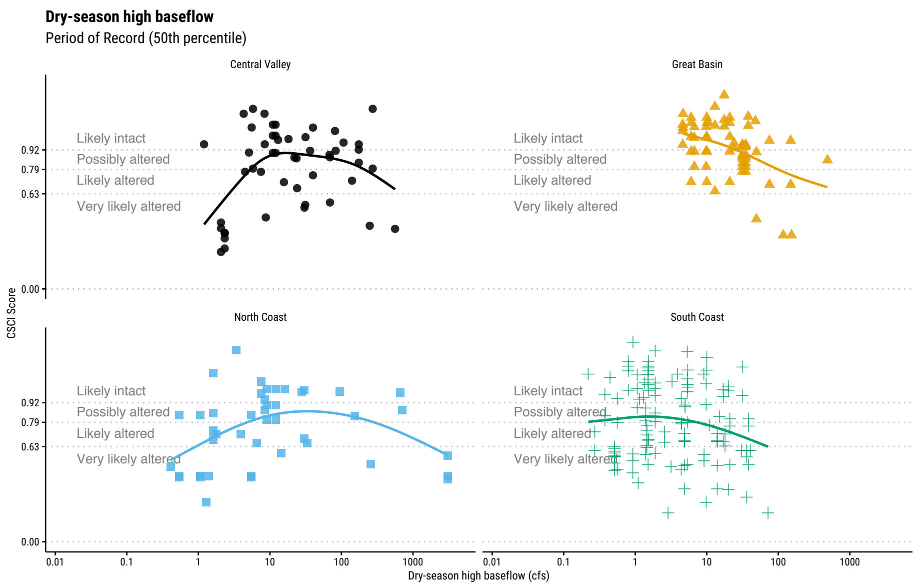

Figure 8. Trend in median dry-season high (90th percentile) baseflow magnitude values relative to CSCI scores for all site pairs in each regional dataset. The trend lines show a general additive model fit to the data indicating that generally for all regional models at moderate discharges, CSCI scores are highest, and at high discharges, CSCI score decreases. Note that the x-axis is plotted on a logarithmic scale, while the y-axis is plotted on an interval scale.

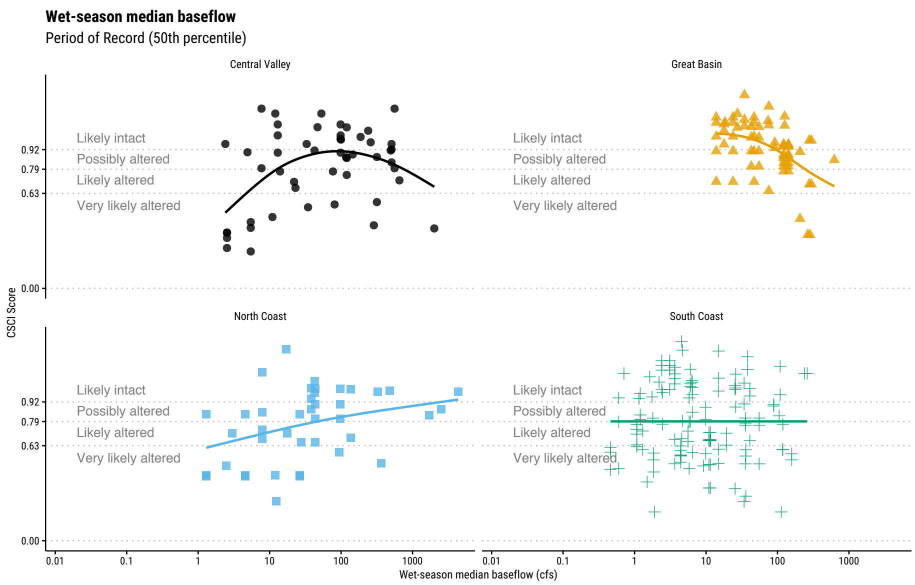

Figure 9. Trend in wet season median baseflow magnitude values relative to CSCI scores for all site pairs in each regional dataset. The trend lines show a general additive model fit to the data indicating that for the Central Valley and Great Basin, at moderate discharges, CSCI scores are highest, and at high discharges, CSCI score decreases. For the North Coast, the trend indicates that as discharge increases, so do CSCI scores. Note that the x-axis is plotted on a logarithmic scale, while the y-axis is plotted on an interval scale.

Application of Results to California Streamflow Management

The results here indicate that broad relationships exist between several key functional flow metrics and biological conditions in streams. In particular, influential functional flow metrics ranged across several functional flow components, indicating that alteration to any of several seasonal flow components (spring recession, wet season and dry season baseflows) may be important in re-structuring biological communities. Further evaluation of these key functional flow metrics with regard to potential thresholds or discrete trends in alteration will help inform the development of ecological flow criteria. For example, decreases in the spring recession rate of change correlated with increasing CSCI score in the statewide dataset and several regional datasets, similar to previous findings in northern Sierra Rivers (Steel et al. 2017). In order to evaluate potential predictive relationships between flow alteration and biological conditions, we will further explore these datasets to assess relationships between the biologic components of CSCI, such as percent clingers or EPT index that are closely related to flow alteration, and functional flow metrics for individual years when sampling occurred.

References

Blodgett, David. 2018. *NhdplusTools: Tools for Accessing and Working with the NHDPlus* (version 0.3.8). Reston, VA: U.S. Geological Survey. https://code.usgs.gov/water/nhdplusTools.

Brown, Larry R., Jason T. May, Andrew C. Rehn, Peter R. Ode, Ian R. Waite, and Jonathan G. Kennen. 2012. "Predicting Biological Condition in Southern California Streams." *Landscape and Urban Planning* 108 (1): 17–27.

De'ath, Glenn. 2007. "Boosted Trees for Ecological Modeling and Prediction." *Ecology* 88 (1): 243–51.

Dormann, Carsten F., Jane Elith, Sven Bacher, Carsten Buchmann, Gudrun Carl, Gabriel Carré, Jaime R. García Marquéz, et al. 2013. "Collinearity: A Review of Methods to Deal with It and a Simulation Study Evaluating Their Performance." *Ecography* 36 (1): 27–46.

Elith, J., J. R. Leathwick, and T. Hastie. 2008. "A Working Guide to Boosted Regression Trees." *The Journal of Animal Ecology* 77 (4): 802–13.

Guisan, A., N. E. Zimmermann, J. Elith, C. H. Graham, S. Phillips, and A. T. Peterson. 2007. "What Matters for Predicting the Occurrences of Trees: Techniques, Data, or Species' Characteristics?" *Ecological Monographs* 77 (4): 615–30.

Lawrence, Justin E., Kevin B. Lunde, Raphael D. Mazor, Leah A. Bêche, Eric P. McElravy, and Vincent H. Resh. 2010. "Long-Term Macroinvertebrate Responses to Climate Change: Implications for Biological Assessment in Mediterranean-Climate Streams." *Journal of the North American Benthological Society* 29 (4): 1424–40.

Lunde, K. B., M. R. Cover, R. D. Mazor, C. A. Sommers, and V. H. Resh. 2013. "Identifying Reference Conditions and Quantifying Biological Variability within Benthic Macroinvertebrate Communities in Perennial and Non-Perennial Northern California Streams." *Environmental Management* 51 (6): 1262–73.

Mazor, R. D., A. C. Rehn, P. R. Ode, M. Engeln, K. C. Schiff, E. D. Stein, D. J. Gillett, D. B. Herbst, and C. P. Hawkins. 2016. "Bioassessment in Complex Environments: Designing an Index for Consistent Meaning in Different Settings." *Freshwater Science* 35 (1): 249–71.

Pebesma, Edzer. 2018. "Simple Features for R: Standardized Support for Spatial Vector Data." *The R Journal* 10 (1): 439.

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 of the United States of America* 104 (14): 5732–37.

R Core Team. 2019. *R: A Language and Environment for Statistical Computing* (version 3.6.2). Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Rehn, Andrew C. 2009. “Benthic Macroinvertebrates as Indicators of Biological Condition below Hydropower Dams on West Slope Sierra Nevada Streams, California, USA.” *River Research and Applications* 25 (2): 208–28.

Ridgeway, Greg. 2015. *Gbm: Generalized Boosted Regression Models* (version R package version 2.1.1). http://CRAN.R-project.org/package-gbm.

Steel, Anna E., Ryan A. Peek, Robert A. Lusardi, and Sarah M. Yarnell. 2017. “Associating Metrics of Hydrologic Variability with Benthic Macroinvertebrate Communities in Regulated and Unregulated Snowmelt-Dominated Rivers.” *Freshwater Biology*. https://doi.org/10.1111/fwb.12994.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2018. “Dplyr: A Grammar of Data Manipulation.” https://CRAN.R-project.org/package=dplyr.

Yarnell, Sarah M., Eric D. Stein, J. Angus Webb, Theodore Grantham, Rob A. Lusardi, Julie Zimmerman, Ryan A. Peek, Belize A. Lane, Jeanette Howard, and Samuel Sandoval-Solis. 2020. “A Functional Flows Approach to Selecting Ecologically Relevant Flow Metrics for Environmental Flow Applications.” *River Research and Applications* 9 (January): 12.