Title: A watershed-specific formula to predict coho salmon reproduction using functional flow metrics

Authors: Kouba, Claire; Jason Wiener; Thomas Harter

General Comments:

The paper addresses the interesting current topic of understanding impacts of flow quantity and timing on native salmonid productivity. The concept behind the study is good, the long timeframe of flow and biological data available at the study site is valuable, and the general idea to statistically assess relationships between flow metrics and fish population size across years is sound. However, I have several major concerns with the conceptual background provided in the introduction and the methodology, which may be impacting the results.

The abstract and practitioner points should be complementary and essentially provide the same information. The introduction lacks mention or discussion of the large body of literature on Functional Flows and how this type of holistic environmental flow assessment approach differs from other flow-ecology assessments. This is very surprising given that the suite of functional flow metrics are included in the methods and analysis.

In the methodology, detail and justification of the flow and fish population metrics chosen is lacking, and the selection of flow metrics to include in the statistical analysis does not address correlations between flow metrics as it should. The selection of statistical modeling method (lasso regression) versus other more commonly used methods needs further justification as well. Due to these methodological concerns, I don't have confidence in the seemingly contradictory results. Some results are complementary to other studies and expectations of salmonid flow requirements, while others are in contradiction with no logical explanation as to why. Further, the discussion does not provide clarifying explanations of either the methodological approach or the results in relation to other studies, leaving me to question the findings.

Given the potential insight that could result from analysis of these long-term datasets, I'd like to see the authors address the concerns summarized here and described in more detail below and resubmit the paper for further review.

Detailed Comments by line number:

Abstract: This is generally clear and well-written. However, what is meant by 'this method for empirically deriving hydrologic metrics' (pg 3 line 14) – were new hydrologic metrics derived in this study or were existing hydrologic metrics used, such as functional flow metrics (e.g. Baruch et al. 2024; Yarnell et al. 2020) or IHA metrics (e.g. Richter et al. 1996). Suggest adding a sentence describing the hydrologic metrics used in the predictive models – more specification is needed due to the hundreds of metrics that have been used to describe streamflow time series in relation to ecological response.

The Practitioner Points should directly relate to what's stated in the Abstract. Suggest adding an additional sentence in the abstract at line 11 (after the sentence beginning "The hydrologic metrics that explain…") along the lines of "Specifically, results indicate that the timing of stream connection and disconnection each year relate to relative coho reproduction and the length of the wet season relates to spawning success." (or whatever is most accurate).

Introduction:

Pg 4 line 40 – The 'functional flows framework' presumably refers to the recent body of literature on the Functional Flows Approach (Yarnell et al. 2015; Grantham et al. 2020; etc.) and the California Environmental Flows Framework (Stein et al. 2021; ceff.ucdavis.edu). Neither Poff et al. 1997 nor Poff and Zimmerman 2010 refer to the term 'functional flows' or a 'framework'. Rather these two papers are foundational for the subsequent work on functional flows, describing natural flow regimes and quantifying ecological responses to altered and unimpaired flow regimes, respectively. It is important for international readers to understand how a 'functional flows approach' and metrics quantifying functional flows are different than other methods to quantify flow impacts on ecological response, such as IHA (Richter et al. 1996) or ELOHA (Poff et al. 2010). A functional flows approach is more holistic (not species specific) and process-based (addresses ecologic, geomorphic, and biogeochemical processes) and, as described in the subsequent sentences, CEFF has been tailored to apply functional flows in California. Please be sure to clarify for readers and cite appropriately.

Pg 5 line 33 – For this sentence "Hydrologic predictors range widely…", suggest including the widely cited IHA metrics (Richter et al. 2016) and ELOHA (Poff et al. 2010) approaches (could cite McManamay et al 2013 here) rather than specific examples of a flow-ecology study. Readers may not be familiar with a specific flow-ecology study, but they are very familiar with various approaches to quantifying hydrologic metrics that are used in these types of studies.

Why is none of the literature on Functional Flows described in the Introduction, particularly given that this is the chosen approach used in this study? On page 6, there is a discussion of what an ideal framework would be to translate relationships between flow and ecological response to water management actions and decisions. This is specifically what the Functional Flows approach and the related CA Eflows Framework describe and do. See https://ceff.ucdavis.edu/publications for literature that should be reviewed and discussed, especially since Functional Flow metrics are used in this study's modeling. The introduction should be updated to include discussion of these concepts and applications, how they are similar or different from the other papers discussed, how these concepts advance the field of flow-ecology and environmental flow science, and what questions currently persist that this study will address.

Methods:

Pg 7 – where does the fish sampling and monitoring (both spawning and juvenile life stages) take place and how often? Please describe and add location(s) to the study map.

Where does the Scott River typically become dewatered in drier years relative to fish monitoring locations? This should be added to the study map (or another figure) to aid in understanding and interpretation of data.

The history of management actions and why in Section 2.2.1 seems out of place here. The management concerns are hard to understand without knowing the differences in the coho versus Chinook life cycle. Suggest moving this to after sections 2.2.2-2.2.4 and shortening 2.2.1 to describe only when, where, and how often the monitoring occurs.

Section 2.2.2 - A key point to mention about the coho life cycle is that they over-summer in streams for one year, which is difficult to do in California's seasonal climate where almost all streams are at very low flow in summer. Hence, coho habitat is often in locations with springs that provide year-round flow or wetlands (likely historically supported by beaver) that retain surface water in the summer (see Lusardi et al. 2020, Lusardi et al. 2021). This directly relates to water management in the valley that reduces already low summer flows to the point of disconnection in more years than not. This also differs from Chinook life histories where fish leave the stream each summer to avoid low flow conditions.

Sect 2.2.4 – An additional factor to mention here is that the location of spawning for coho versus chinook also likely has to do with requirements for where water remains in the summer. Coho spawn in the tributaries not only because of sediment size, but that's the only place where water remains in the summer for rearing juveniles. Chinook can spawn in the mainstem because their smolts outmigrate before the river dries out.

What are the 88 metrics and how is this similar or different from 'an initial set of metrics' from the functional flow metrics cited? The number of functional flow metrics in Yarnell et al. 2020/Patterson et al. 2020 relating to each of 5 functional flow components is 24. Table 1 lists many metrics that are not included in the cited functional flow metrics, including river reconnection/disconnection at a variety of thresholds, minimum flow in some months, total annual flow, and days above certain flow percentiles. Why were these metrics chosen versus other any other metrics that relate similar aspects (e.g. from IHA). The functional flow metrics and related citations help to provide an explanation of why those 24 metrics are used, which should be briefly summarized for readers unfamiliar with those papers here. Additionally, explanation of why the other metrics were selected along with any applicable references should be included.

Pg 15 line 10 – please provide a discrete definition of these connectivity metrics – For each year, this is the date (presumably day of water year) that a certain flow threshold is met? If the stream stays wetted all year are these metrics a null value? Table 1 shows multiple flow thresholds and states this metric is the number of days after Aug 31, which implies a duration not tied to a day of water year. Please clarify. Additionally, how do these dates relate to the timing of the spring recession start and full pulse timing? Their calculation appears to be independent.

Pg 15 line 16 – The 2-year flood flow (wet season peak flow 2-year; calculated as the 50th percentile of peak flows over the period of record) in the suite of functional flow metrics approximates scouring flows in stream channels based on geomorphology literature. Why is the 90th percentile flow also chosen to represent scouring flows? How is this different than the 2-year, 5-year, and 10-year peak flows calculated? Often in fish studies, there is an estimate of the flows needed to move spawning gravels – it is likely more applicable to just use that mobility threshold as an additional metric than the 90th percentile flow.

Pg 15 lines 32-35 – please explain what these two sentences mean – how is the timing correlated? What types of correlations are expected? Details about the statistical analysis and why one method was chosen over another should be discussed in a separate section.

Figure 5 is helpful to explain the connectivity metrics. A similar figure for min-flow and total flow metrics in the relevant periods would be helpful. In the caption, why is 'more detail on selection of flow thresholds' in the Results section –this should be in the methods section.

Generally, section 3.1 needs more detail, justification of metric choices and better organization stepping through all metrics selected in succession.

Sect 3.2 - What are the methods for collecting this population data? Who collected and analyzed this data if it wasn't within this study? Presumably this data was collected by other scientists (agencies, local watershed groups?) and this study is assessing that data. If this is the case, then the reports or publications that fully explain the methods, data analysis, and results for all metrics need to be properly cited and summarized here.

Sect 3.3 –Which table is Supplemental Table 2? The supplemental tables (both 1 and 2) are not labeled or formatted well for readability. It's unclear in the proof which tables are which, column headers are cut off, and many cells are simply X's. These tables should be clearly formatted with explanatory captions, repeating headers for each page, and properly sized cells so the data is readable.

Sect 3.3.1 – The concept of a brood year, rearing year, and smolt year is very confusing as it doesn't overlap with water year and the brood year and smolt year are not a full year. There is no reference/citation to other studies that use this terminology, and this is not commonly done in studies I'm familiar with. In reality, these are life stages of coho that have differing durations and occur over different seasons. For clarity and accuracy, it would be better to just refer to these time periods as 'smolt stage' and 'brood stage' or 'smolt period'. Fall pulse flows in any year then affect the brood stage and the rearing stage, while the spring recession affects the rearing and smolt stages.

Pg 21 line 6, 13-14, and subsequent paragraphs – This is all very confusing. Presumably, a February 2014 flow reconnection would be associated with Water Year 2014 and could affect the 2013 brood stage, 2014 rearing stage, and 2014 smolt stage – if life stages were defined by calendar year. If life stages were defined by water year (which would better overlap with the flow data), then a Feb 2014 reconnection date affects the WY 2014 brood stage, WY 2014 smolt stage, and WY 2014-2015 rearing stage. Because the rearing stage overlaps water years, you could specify a WY 2014 early rearing stage and a WY 2015 late rearing stage. Defining all metrics based on WY will help align the statistical analysis and limit confusion on what is meant by 'year'.

Similarly, a 'brood year' is defined as a specific stage (4 month spawning window) but then also referred to as a cohort (line 17). Suggest using very clear simple defined terms consistently. For example, the 'WY 2010 cohort' would begin in WY 2010 and end in WY 2011, including the 'WY 2010 brood stage', 'WY 2010 early rearing stage', 'WY 2011 late rearing stage', and 'WY 2011 smolt stage'. This would eliminate the need to then adjust 'years' as described in lines 21-23 and make for much easier to understand metrics and terminology.

In figure 3 and previous text, the coho spawning stage is described as commonly occurring from October to January, but 'brood stage' is defined as starting in September. How often did spawning actually occur in September in the dataset? To help overlap with water year, it would be easiest to just define 'brood stage' as Oct-Jan.

Sect 3.3, like section 3.2, needs more clarity on how the fish response metrics (i.e. the number of each life stage of fish in each water year) were evaluated against the flow metrics (which are defined by water year). Presumably the statistical model evaluates which flow metrics (predictors) best relate to each fish metric (response), and the entire dataset can be evaluated simultaneously. Does the chosen statistical model require that only a subset of the flow metrics be compared to each fish response variable? If so, then this should be clarified – e.g. "Each cohort life stage was compared to the flow conditions in the associated water years they experienced". Or is there another reason why this relation between flow metrics and fish life stage metrics needs to be defined?

Section 3.4 – why is the correlation analysis done on predictor-response pairs versus just the predictor variables as is standard? It may be that many of the flow predictor metrics are correlated (e.g. spring recession timing and spring recession magnitude, or wet season timing and flow reconnection date), but presumably the statistical analysis will determine which flow metrics are correlated or not with the fish response metrics.

Section 3.5 - This section doesn't seem to be needed.

Sect 3.6 – Why was lasso regression selected as the statistical analysis method versus other more commonly used methods such as MARSS or Random Forest modeling? Please include a justification for this method.

The methods sections could be condensed into sect 3.1 – explaining the flow predictors used in the statistical analysis, sect 3.2 – explaining the fish response variables used in the statistical analysis (all life stages for each species or just one response variable for each species representing the productivity of each cohort e.g. # smolts per spawner or both), and 3.3 – explaining the statistical analysis, whereby either a single model was determined for each species (a model describing the flow metrics that best predict # smolts per spawner) or several models for each life stage of each species (each describing the flow metrics predicting # of fish in that stage).

Sect 3.6.2 - If lasso regression can't deal with null values in the dataset (e.g. some years with no fall pulse calculated), and some years don't have fish data for some life stages, what years ended up being excluded?

Sect 3.7 – In essence, Lambda was then 0 to result in this function? And HBi is equal to yi, the observed value of the ecological response. If this is the case, then section 3.7 isn't really a next step from section 3.6. They could be combined into a single statistical analysis methods section describing the statistical model and how it was adjusted to fit the dataset (via lambda).

Results:

Sect 4.1, figure 6 – Why show this subset of FF metrics? This could show only those metrics that have obvious increases/decreases over the period of record. I would guess all the magnitudes have changed, along with some timings and durations. Please ensure the names of these different metrics are the same as in the literature cited and at ceff.ucdavis.edu.

Sect 4.2 – A correlation between a single flow metric and a single life stage response variable isn't descriptive or causal in explaining the impacts of flow on life stage given all the other interrelated factors (as evidenced by the low correlation values). Multi-parameter statistical models will be more meaningful, allowing for relationships between different predictor variables to be considered. The correlation analysis is most useful for understanding correlations between predictor variables to allow for selection of those predictors that are more independent and empirically make sense with regard to the response variables. Was this analysis used to select certain variables to include in the modeling? If so, then that needs to be clarified and results should focus on that point. Inferring effects of hydrology on fish response metrics from a simple correlation is not appropriate within this context.

Pg lines 47-51 – Perhaps this is an indication that lasso regression is an inappropriate method for anlayzing this data. Other flow-ecology studies have used random forest models (e.g. Peek et al. 2022; Steel et al. 2017), multivariate autoregressive state-space (MARSS) models (e.g. Baruch et al. 2024), and boosted regression tree analysis and generalized linear models (GLMs) (Irving et al. 2022; Sengupta et al. 2018) to assess relationships between flow metrics and ecological responses. If the authors are confident in their approach, justification for this approach versus other commonly used statistical methods should be included in the methods section (see previous comment).

Pg 33, line 9-10 – If there's a great deal of collinearity between several metrics (which there are), then how can you know which metrics are actually relevant? This is what a pearson's correlation can address – it's a method to limit the number of predictor variables to those that are least correlated but ecologically significant. Similarly, is it really the spring recession rate that's informative for coho or is it the timing of spring disconnection or is it the duration of the spring recession? I would expect these 3 variables to be highly correlated but with different ecological interpretations (slow recession rates mean slow changes in river stage and velocity, which can have a different impact on outmigration than the timing of disconnection where passage is no longer viable) and different implications for management.

Next paragraph – what is "log-transformed total flow in the brood year fall"? Is this the same as total flow listed in Table 1? Of the variables listed here, I would expect many of these to be correlated with each other. The conclusion that coho may be more sensitive to timing metrics, while Chinook may be more sensitive to magnitude metrics may be correct; however, the statistical models don't conclusively show this given the highly correlated nature of many of the flow metrics.

Tables 5 and 6 suggest that higher coho numbers are associated with lower values of DS_Mag_90 and higher chinook numbers are associated with lower values of sep-dec flow. This is counter-intuitive and contradictory to earlier reconnection and later disconnection timings. Longer wet seasons occur when flows are higher.

Sect 4.4.1 – The magnitude predictors for coho and Chinook are counter-intuitive and don't match what is generally known for fish habitat. The timing predictors for coho, which do make sense, don't sync with the magnitude predictor. I wonder if the correlations between flow metrics are muddled in the statistical analysis and obscuring findings.

Sect 4.4.2 – Using the key predictors, determined from observed values, to predict those same observed values isn't useful to the analysis. They simply show that the predictors are related to response variables, which the statistical analysis shows.

Sect 4.4.3 – similar to sect 4.4.2, this section is just restating what the statistical analysis showed, that certain predictors relate to production. It doesn't add any new or additional information.

Discussion:

Sect 5.1 – This would be better framed as a discussion around the utility of a statistical approach using flow metrics that quantify ecologically important aspects of the flow regime – determined by using functional flow metrics and reconnection metrics that are not correlated. Comparing this approach to other flow-ecology approaches lends justification to this approach. See Baruch et al. 2024 for how this concept could be discussed and framed.

Sect 5.2 – 1st paragraph - how are scouring flows a finding? None of the peak metrics were significant and gravel mobility flows weren't assessed. Line 20 – what metric represents 'scouring days'?

The 3rd paragraph is well written, but the second sentence (lines 27-28) seems contradictory to the conclusion in the rest of the paragraph.

Sect 5.3 – determining flow metrics that relate to connectivity appear to be important for this system and are not part of the suite of FF metrics. However, not addressing the correlations among metrics lends to ambiguity of which metrics are actually relating to fish responses. Is it the date/timing of flow connection or the magnitude of flow connection that is the ecologically important aspect? Is there a known magnitude based on other studies that provides connectivity to 'good' habitat versus just 'any' habitat? Using correlation analysis to determine which connectivity metrics are least correlated and selecting among the highly correlated variables those that are ecologically meaningful (e.g. a low magnitude connectivity threshold that provides any passage (40 cfs) and a higher magnitude threshold that provides connectivity to 'good' habitat (100 or 200 cfs)) will provide a stronger case for the statistical results.

Sect 5.4 – I'm not convinced that the study results support a conclusion that Chinook are less sensitive to flow conditions than coho. Given the correlations among flow metrics and the variety of other flow-related factors (e.g. temperature) that may affect salmon survival, the statistical results are not that conclusive. Rerunning the models with a paired down set of flow metrics that are less correlated and more directly related to ecological outcomes may yield different results.

The bulleted points here are only partially correct – see previous comments regarding substrate preferences, habitat density issues (internal population dynamics) are typically associated with limited habitat availability at lower flows, differences in ocean conditions affect both species and are not examined and discussed in this study.

Sect 5.5 – Using the statistical model as a predictive tool of fish productivity isn't the correct use of this analysis. There is not a 'direct effect' of a particular flow metric or set of metrics on fish response (line 32) – just a positive or negative statistical correlation between the variables. Most flow-ecology analyses seek to understand which flow metrics correlate to an ecological response in order to understand thresholds around which flow management may impact or support desired responses. To actually predict fish population responses, more factors affecting fish productivity need to be included (e.g. temperature, water quality, etc.) and detailed understanding of the various life stages need to be incorporated (e.g. as in individual-based models).

Discussion of how exactly the study findings might inform management would useful. Are there actions that might support increasing the wet season duration and extending the spring recession duration?

Conclusions:

The idea of hypotheses to be explored based on study results is good, but not discussed in previous sections. These could be a way to expand the discussion past the statistical findings and into considerations for management.

Pg 44 line 12-13 – Is the main conclusion that flow during the spawning period is most related to fish production or is that longer wet seasons, affecting both the spawning and early rearing periods, are related to greater fish production?

The first sentence of the last paragraph implies that climate change is the reason for shorter wet seasons, with little options for managers to improve conditions. The declining trends were explained in the introduction as due to increased consumptive use and groundwater withdrawal. Management actions that limit these factors may improve conditions in moderate and wetter years, as shown in some of the related groundwater modeling studies (which could be referenced here in discussion).
