Independent Scientific Advisory Board
for the Northwest Power and Conservation Council,
Columbia River Basin Indian Tribes,
and National Marine Fisheries Service
851 SW 6th Avenue, Suite 1100
Portland, Oregon 97204

Review of the Comparative Survival Study (CSS) Draft 2022 Annual Report

Members

Courtney Carothers
John Epifanio
Stanley Gregory
Dana Infante
James Irvine
Yolanda Morbey
Thomas Quinn
Kenneth Rose
Carl Schwarz, Ad Hoc
Desiree Tullos
Thomas Wainwright
Ellen Wohl

ISAB 2022-1
October 18, 2022

ISAB Review of the Comparative Survival Study (CSS) Draft 2022 Annual Report

Contents

ISAB Review of the Comparative Survival Study (CSS) Draft 2022 Annual Report

I. Background

The Columbia River Basin Fish and Wildlife Program calls for a regular system of independent and timely science reviews of the Fish Passage Center's (FPC) analytical products. These reviews include evaluations of the Comparative Survival Study's draft annual reports. The ISAB has reviewed these reports annually beginning twelve years ago with the evaluation of the CSS's draft 2010 Annual Report and most recently the draft 2021 Annual Report.[1] This document is the ISAB's review of the *draft 2022 CSS Annual Report: Comparative Survival Study of PIT-tagged Spring/Summer/Fall Chinook, Summer Steelhead, and Sockeye* and is the ISAB's thirteenth review of CSS annual reports.

II. Summary

This ISAB review begins with an overview of the latest report's findings (this section). It moves on to suggested topics for further CSS review (Section III), and then general comments and specific editorial comments on each chapter of the draft 2022 CSS Annual Report (Section IV).

The annual CSS report is a mature product, typically including mostly updates with the latest year of data and expansion of analyses as more data are acquired. Many of the methods have been reviewed in previous ISAB reports and so now receive only a confirmatory examination. As more data are acquired, new patterns and questions arise on the interpretation of the results—this is now the primary focus of our reviews. The ISAB appreciates the CSS's detailed responses to suggestions provided in previous reviews (e.g., CSS 2021 Annual Report, Appendix H), and we do not expect the CSS necessarily to act immediately to new requests for further analyses.

The Fish Passage Center has produced these reports since 1998, and the ISAB has reviewed them since 2010. As a result, much of the CSS report (Chapters 1, 3, 4, and 5) focuses on the continuation of the analysis of long-term trends. Much of the text in the reports is taken

[1] ISAB 2010-5, ISAB 2011-5, ISAB 2012-7, ISAB 2013-4, ISAB 2014-5, ISAB 2015-2, ISAB 2016-2, ISAB 2017-2, ISAB 2018-4, ISAB 2019-2, review of Chapter 2 of the 2019 Annual Report (ISAB 2020-1), ISAB 2020-2, and ISAB 2021-5.

verbatim from the text of previous reports, with changes to update the time periods and results with additional data. Averaged values calculated over time series change very little because the additional year of data represents a small fraction of the total record. For most chapters, the final conclusions are almost identical to conclusions in previous reports.

The ISAB previously encouraged the CSS to describe major applications of the CSS data that have been published or reported over the last year and briefly highlight important findings based on CSS data. The ISAB continues to recommend that the CSS should make the regional importance of its analyses abundantly clear to the Council, BPA, co-managers, and the public.

In 2021, the ISAB noted that many things have changed in the Columbia River Basin and hydrosystem over the 25 years of data collection, and the effects of these changes on the long-term analyses of salmonid survival are largely unknown. We suggested a description of important changes in the hydrosystem over the years and a brief indication of their possible effects on salmonid survival. In response, the CSS developed a new chapter (Chapter 7) for the 2022 Annual Report that provides a thorough and readable account of the history of the hydrosystem, development of management of spill to benefit salmon and steelhead, and the formation and evolution of the CSS study to inform state, federal, and tribal fisheries managers. The ISAB appreciates the CSS response and is confident it will be informative for the region.

Regarding Chapter 2, the CSS Oversight Committee developed an Adaptive Management Framework as a response to proposals for Spill Experiments. The Adaptive Management Framework calls for measuring several key metrics on an annual basis: juvenile Fish Travel Times (FTT), juvenile survival, ocean survival, smolt-to-adult survival rate (SAR), and transport to in-river smolt-to-adult survival rate (TIR). The Adaptive Management Framework also acknowledges several drivers that influence the performance of the key metrics for salmon and steelhead under the Flexible Spill Agreement, including water travel time (WTT), expected number of powerhouse passage experiences (PITPH), ordinal day, and total dissolved gas (TGD). As the Preferred Action (PA), the US Army Corps of Engineers (USACE) began operating the dams on an 8/16-hour schedule of low/high spill ratio since 2019.

This is the first evaluation of the Flex Spill experiment that appears in the CSS annual report. Because the complete return of cohorts tagged during the experiment are not yet available, only the Fish Transit Time (FTT) and in-river survival are evaluated. The results from the Flex Spill experiment for these variables appear consistent with modelled values based on 1998-2018 data, but no formal assessment is done. The analyses to date indicate that spill levels that result in up to 125% TDG, or perhaps even greater, would not decrease juvenile survival. Doing such analyses now, even with incomplete data, is highly encouraged, and the ISAB offers suggestions for additional improvements in data analysis and the use of results to inform the adaptive management process.

Chapter 3 continues and expands previous years' work on the effects of the in-river environment on juvenile travel time and survival. The CSS now analyzes the Upper Columbia River based on PIT tag data from Rocky Reach to McNary dams rather than Rock Island to McNary dams.

Chapter 4 also continues and extends past years' work on patterns in annual overall SARs. By now, the low level of SARs relative to the Council's 2%-6% objectives has been established. The CSS could consider developing an Impact Report, perhaps developed collectively with other groups, to communicate the most critical take-home messages and implications for the Council, BPA, and co-managers.

Chapter 5 presents analysis of upstream migration success for adult spring and summer Chinook, fall Chinook, and steelhead. All survival probability estimates are very high, with two segments having survival estimates very close to 1.0. The ISAB is concerned that the lack of contrast in survival in many reaches over time will make it difficult to determine effects of other factors. What is the end-goal of this analysis? What are the major management implications from these results? This chapter is close to becoming a "mature" chapter with only minor updates each year as new data are collected and added to past data.

Regarding Chapter 6, the premise and motivation for this new chapter make it a great addition to the annual report. The link between SARs and abundance targets is fundamental to understanding what restoration and other recovery-related actions must achieve to be successful. The analyses are designed to address the question: "What level of SAR is needed to realize the Columbia Basin Partnership (CBP) benchmarks, and is that level of SAR consistent with the prescribed NPCC goals?" In addressing this question, the analyses also provide a synthetic view of the status of the situation for an impaired population (i.e., Snake River spring/summer Chinook salmon).

In Chapter 6, the data are examined with a multiple model analysis designed to increase confidence (robustness) when different models (approaches) agree. However, without careful examination of the assumptions (e.g., in common, alternative, unique) and the details of how common data are used across the models, it is impossible to assess the degree to which the predictions from the three approaches are independent. It is clear that the three models are not independent (e.g., as shown below, approaches 2 and 3 are very similar). The assumptions and sources for key inputs, especially for approaches 2 and 3, are insufficiently documented to assess their commonalities. Approaches 2 and 3 might be better presented as alternatives of the same approach.

Unfortunately, this first presentation of Chapter 6 reads as though the three groups each did an analysis and then they linked together their written reports describing their methods and results. Consequently, the presentation of the three approaches is confusing because: (1) the methods and results were not merged and integrated sufficiently (the Discussion does some of this), (2) approaches 2 and 3 seem very similar as they both end with the same estimating equations (except for the conversion factor C) but appear as separate analyses (see detailed comments below), (3) the input data and information used have unknown overlap, especially between approaches 2 and 3, and (4) the empirical model for approach 1 could be improved.

The current version of Chapter 6 needs additional work to resolve the issues of concern mentioned above and in our detailed review below. Nevertheless, the main message is that substantially larger SARs than the 2%-6% recommended ranges from the CSS reports will be needed to meet the recovery goals. The need to address key management questions should be highlighted in a Conclusions section for the Council and regional decision makers. The many years of CSS reporting should provide valuable information on how to prioritize management actions.

Chapter 7 summarizes how spill has changed since the completion of the hydroelectric projects in the early 1980s. The ISAB appreciates the difficulty in summarizing spill changes that have occurred on many different temporal scales, and we found this high-level summary very informative. In particular, the following conclusion should be given more prominence because of the many conflicting arguments about the need to change spill over the years:

> "*The impact of changes in spill on metrics such as juvenile survival and SAR are difficult to discern in the time series because dam passage route is only one of many processes that influence survival. Said another way, relatively small changes in spill occurred over a period with high variation in river discharge and highly variable ocean conditions*."

Importance of Information in CSS Reports

As we stated in our review of the 2021 CSS Report, the ISAB strongly emphasizes the importance of the CSS reports. There may be a tendency to think that the annual CSS reports are "more of the same" each year; however, with more than 25 years of data, the conclusions reached are now at the stage where the uncertainty in the results are well estimated and often small (and not just artefacts of particular years). As well, the physical (e.g., PIT tag detection arrays) and human capital infrastructure (e.g., logistical capacity and expertise) added over the 25 years is extremely valuable going forward. Long-term records of fish abundance and environmental conditions are extremely difficult and expensive to develop. Survival of salmon and steelhead during their life cycle is significantly affected by the hydrosystem, and these data are essential for the Fish and Wildlife Program. The ISAB encourages the Council and BPA to reassess the funding and consequences of flat funding since 2017. This funding policy directly affects the quality of the research and integrity of the long-term records. If funding is to be constrained, then the entire CSS program should be reviewed in detail to assess which parts are essential and which parts can be reduced without critically damaging the overall CSS program.

III. Suggested Topics for Further Review

Since 2011, the ISAB has suggested topics that warrant further CSS or regional review; see Section V below for the ISAB's evolving lists of topics. The latest CSS report incorporates many of our past suggestions. As noted above, the ISAB appreciates the CSS's effort to respond to our past comments and suggestions.

Some of the past recommendations from the ISAB appear to be beyond the current scope of the CSS (see several from 2017 in Section V) but will become increasingly important in the future. Some of our earlier and current recommendations may seem repetitive and unachievable within a year to inform the next report, but they deserve some advance planning as these issues will become much more pressing in the future. In particular, any data gaps should be identified for potential new data collection procedures. Similarly, any modifications of life-cycle models should be flexible enough to incorporate these issues. This is reflected in our recommendations for future work below.

1. Given that the Council's SAR targets are generally not being met, this could imply that the populations are more or less destined for functional extirpation sometime in the future. Chapter 6 identifies SARs needed to attain the higher goals of the Columbia Basin Partnership, and their analysis indicates that "higher SARs than seen in most recent years are needed to attain and maintain higher natural-origin abundances, rather than mere persistence." Factors related to attaining the recommended SARs need to be explained with respect to the suite of actions implemented under the Fish and Wildlife Program.

2. Although the CSS is an empirical modeling effort, can the FPC and CSS Oversight Committee expand upon previous analyses to identify further evaluation and data needed to address the "breaching" proposals for the four lower Snake River dams more fully? This has been a critical regional issue for more than two decades and is currently being discussed in state and federal documents (NMFS 2020, Columbia Basin Bulletin 2020, 2021, 2022, NMFS 2022; also see Storch et al. 2022). Insights provided by the depth and scope of CSS analyses over the last 27 years will be important in these regional discussions. Chapter 2 of the 2019 CSS Annual Report included breaching as one of the EIS alternatives considered and found that it resulted in the highest SARs of the alternatives examined. Is breaching an all or nothing proposition, or can significant gains be expected with fewer dams being breached?

Additional analysis and interpretation by the CSS would be informative and would strengthen the use of the CSS data in future analyses by others.

Status of CSS responses and progress on ISAB suggestions from 2021 review

In ISAB 2021-5, we recommended the following six topics for future reports. After each recommendation, we summarize the current status of the work to address them:

1. Provide more robust introduction section that includes a summary of major findings, highlights new analyses, and describes recommendations for potential management applications of findings. Describe changes in annual report structure from year to year, including why chapters and analyses were dropped or added.

ISAB Comment 2022: No changes were made in the 2022 report in response to this ISAB comment. We appreciate that subjects are introduced and then resolved and so chapters in the CSS are added or removed. Perhaps a table showing chapter contents by year and an explanation of why a chapter is dropped would be a nice summary to help the reader navigate the many years of the CSS. This could be an appendix that is updated each year.

2. Describe major applications of the CSS data that have been published or reported over the last few years and briefly highlight the important findings that are based on CSS data.

ISAB Comment 2022: The CSS improved language as needed and updated references. The ISAB encourages continued diligence in keeping the annual reports up-to-date by citing and incorporating results of other studies and literature that provide relevant information that inform CSS analyses or use CSS products.

3. The CSS is pulled in many directions (and sometimes sent along a new path by the ISAB). Given the funding constraints identified earlier, it will be useful to step back, decide on the core results that need to be presented, identify the major uncertainties in the results and how these could be addressed. Which results are now of lesser relevance? Which results will be crucial for the future in the face of climate change, habitat restoration effect, and hydrosystem management changes?

ISAB Comment 2022: The CSS agrees that they simply respond to the agenda as set by the co-operating agencies and they are lacking in resources. Nevertheless, we reiterate the need for a longer-term vision to help allocate resources for the future CSS reports. Perhaps the CSS reports should be split into a “core” document and then additional chapters to address the various concerns as they arise?

4. Explore analytical methods to adjust for biases for smolts captured and tagged at Rock Island to maintain a longer period of information.

ISAB Comment 2022: In the 2022 draft annual report, the CSS examined paired observations for both yearling Chinook salmon and steelhead and found that although both species showed evidence of bias, the magnitude of the bias varied across the cohorts. They concluded that applying a constant correction factor across historical cohorts would not be appropriate because of this variation in the magnitude of the bias. This analysis deserves further investigation. Are there cohort-specific covariates available which might predict the magnitude and size of the bias by cohort?

5. Address the unusually high mortality rates of subyearling Chinook in the MCN-BON reach and include major recommendations in their Conclusions.

ISAB Comment 2022: This issue is being investigated by the CSS authors, but the information available currently is not sufficient to include conclusions in the 2022 report. The ISAB has repeated this recommendation.

6. Form a working group to explore how newer computer technology could reduce the human cost of updating and reporting the CSS report.

ISAB Comment 2022: The CSS feels that their process is already highly automated and are not considering any further major changes in workflow. With the new chapters added, especially Chapter 6 with new analyses, the ISAB encourages the CSS to look for efficiency gains in data preparation and analyses.

IV. Comments on New or Updated Analyses in the draft CSS 2022 Annual Report by Chapter

IV.A. Comments on the Executive Summary and Chapter 1. Introduction

The Executive Summary highlights the major findings for each of the chapters, which is consistent with previous ISAB recommendations for the CSS Report. One of the serious challenges is the loss of financial support caused by BPA's flat funding policy from 2017 to the present. The ISAB has emphasized the importance of the CSS study in past reviews and raised concerns over the erosion of funding to provide this critical information. The ability of the CSS to respond to ISAB reviews and technical advances is severely constrained by the flat funding effectively resulting in decreased financial support over time.

The 2022 CSS Report includes three new chapters. Chapter 2 describes the development of the Adaptive Management Framework for evaluating hydrosystem operations and the application of CSS studies in the adaptive management processes and past decisions. Chapter 6 provides an analysis of the smolt-to-adult returns (SARs) that would be required to meet the low, medium, and high goals of the Columbia Basin Partnership Task Force (CBP). This more quantitative and technical chapter describes three approaches (models) for determining the numerical goals. Chapter 7 responds directly to previous ISAB recommendations and describes the history of the hydrosystem, development of management of spill to benefit salmon and steelhead, and the formation and evolution of the CSS study to inform state, federal, and tribal fisheries managers.

The three new chapters are valuable additions to the CSS Report. The Report does not indicate whether the chapters from the 2021 Report that were omitted in 2022 will be included in future reports:

1. Patterns of Survival of Wild Spring Chinook Salmon and Steelhead in the Columbia River Basin
2. Development and Assessment of an Approach to Estimate Daily Detection Probability and Total Passage of Spring-Migrant Yearling Chinook Salmon at Bonneville Dam

A brief description of the plans for these studies would be useful for fisheries managers, the Council, and readers.

Chapter 1. Introduction is an update of this chapter from previous reports. As we have commented before, this chapter provides a comprehensive and reasonably readable overview of a complex program. There is little change other than addition of numerical results from 2021 and information about additional sources of data. The CSS incorporated many of the changes the ISAB recommended in 2021.

IV.B. Comments on Chapter 2. An Evaluation of Adaptive Management Changes in Hydrosystem Operations on Salmon and Steelhead Survival and Travel Time

This chapter is the first analysis of the results of the Flex Spill experiment. Because not all fish have returned from the cohorts affected (2019-2021 brood years), only the Fish Transit Time (FTT) and in-river survival of the affected cohorts are examined.

The strategy used in this analysis is to (a) fit a model for the FTT and in-river survival using the 1998-2018 (pre-experimental conditions), (b) use the fitted model to predict the FTT and in-river survival using the conditions present during the Flex Spill experiment, and (c) compare the FTT and in-river survival from the Flex Spill years estimated from field data with the model-predicted values.

No formal statistical assessment of the degree of model-data agreement is performed, and the evaluation of the fit of the latest 3 years is mostly visual to see if the data-based estimates seem to match the predicted values and to see if the modelled values fall within a confidence interval for the estimates for the newer data.

The premise and idea of the chapter are excellent, and the ISAB encourages the CSS to continue such analyses. Conducting analyses as soon as the data become available is recommended, as that will allow the analysts to improve the process in advance of further data. In its present form, the analyses show promise but could be improved in the "learning" part of Adaptive Management. Now is a good time to show and demonstrate how the analyses can help with the ultimate goal of informing that adaptive process.

Areas that should be addressed to improve the analyses are: (1) statistical approach, (2) use of explanatory variables, (3) presentation of results, and (4) linkage to adaptive management to inform future actions.

Substantive comments:

Statistical approaches:

A more formal assessment of model performance could be conducted by making "leave-one-out predictions" for the initial data series and finding the distribution of a discrepancy measure for these "leave-one-out predictions" and then comparing the actual discrepancy in the last 3 years to this distribution.

The predictions of FTT and in-river survival from the fitted model applied to the Flex Spill years also have uncertainty, which is not presented in any of the plots. It may be difficult to determine an appropriate measure of uncertainty because the predictions are "marginal" (i.e., ignoring random effects).

An alternate analysis would be a Bayesian approach, because all sources of uncertainty (e.g., random effects) can be accounted for in the credible/prediction intervals. Similarly, a Bayesian p-value approach may be easier to implement to assess if the newer data are congruent with the model based on previous data.

While comparing predictions of biological metrics to observed data, there is much to learn by comparing model predictions under various conditions (real and targeted) as well. How this statistical approach can interact with other approaches (e.g., LCM, SAR estimation) should be a part of the analyses to support AM. The analysis in this year's report is presented as a stand-alone analysis. This is understandable from the point of view of doing and describing analyses but more information can be obtained by combining this with other models to achieve the ultimate goal of informing the adaptive management process. Is that the task of another effort or another group?

Use of explanatory variables:

It appears that the authors used the observed (or field-estimated) values for explanatory variables (PITPH, WTT) for 2019-2021. The predictions obtained that use these explanatory variables are thus a prediction of the biological metrics we would expect to see in nature. If there are differences between the observed and explanatory variables for the Flex Spill experiment, then predicting the biological metrics for those planned conditions would be very helpful in the interpretation. Was the small response due to the Flex Spill not having the expected effects on PITP and WTT or was it due to the biological response to PITP and WTT being muted? If the Flex Spill goals were achieved, what would be the biological response? The chapter refers to these types of analyses being done in planning of the Flex Spill. It would be useful to present them here to ensure direct comparability to other results in this chapter. The analyses would suggest whether the actions should be refined or modified so the desired effects on explanatory variables are realized or that the actions should be altered to better reflect the biology.

A more straightforward presentation of the time series of the explanatory variables, such as a stacked set of plots, would be informative. In addition, maybe the model does not adequately control (e.g., missing variables) for other conditions that may be influencing the biological responses. Examining the conditions of 2019-2021 relative to historical conditions would help with interpretation now and when more years of data become available. The idea would be to identify whether there were unusual values of factors, other than those included in the model during 2019-2021, that could “mask” or affect the biological responses to Flex Spill.

While more data are needed, “testing” analyses for their ability to answer these types of questions now will prepare the analysis for their application when more data become available. If time permits, there are sufficient data to conduct simulated data testing of the methods. Understanding the power of the various statistical analyses will become critical in future analyses. Earlier analyses predicted small responses.

Presentation of results:

The predictions for 1999-2021 should be part of the historical plots to allow the reader to see differences before and after the Preferred Alternative.

The presentation of results for the TDG is confusing. In particular, positive coefficients meant that high TDG was associated with higher survival. The solution, which seems sound, was to use residuals after adjusting for WTT and look for a trend of them declining with increasing TDG.

The explanation presented for the negative correlation between WTT and TDG and for the solution used is complicated and difficult to follow.

Linkage to Adaptive Management to inform future actions:

While it is early to generate results to actually inform adaptive management, the CSS should determine how such results could be used. The Discussion section in this chapter begins to show potential management options going forward (e.g., last paragraph on page 82), but the link back to the actual analyses is vague and the future link to adaptive management is also vague (necessarily at this time). A mock analysis is suggested to see if and how the results can be used, without presenting the results as actual conclusions. The chapter is focused on data → statistical modeling → adaptive management, with the idea to document observed biological responses. At some point soon (maybe not in this chapter this year), the authors should consider how to do the analyses and present their results from the perspective of adaptive management → statistical modeling → data. For example, using the same modeling as reported in this year's report, one could explore how different combinations of explanatory variables affect the biological metrics and whether any of those combinations resemble the targets of the Flex Spill experiment and how do "successful" combinations overlap and differ in their changes to explanatory variables. This would be combined with the field data-based analyses reported this year and power analysis and other uses of simulated data.

Minor comments:

p. 62. Ordinal day is used to account for seasonality in the models. It would be helpful to know which factors of seasonality are meant to be accounted for by this measure, along with evidence that the factors are not likely to change with climate. If such evidence does not exist, it could be useful to integrate additional measures into the models going forward.

p. 65. The authors state, "the PITPH index may or may not reflect the actual proportion of smolts that pass through the powerhouses," indicating that this is an outcome of the Flex Spill operations. The ISAB appreciates the illumination of this point but does not find any specific strategy to address the point. How is it accounted for in the analytical models? Because of this, it could be helpful to integrate that uncertainty into the modeling framework in some way.

p. 65. "Compared to previous years, … comparable to other cohorts in years 2001, 2010, 2015 (Figure 2.3)." The reason is probably the drought conditions in that part of the Columbia River Basin, but this should be confirmed if it is drought related or explained if it is not drought related.

p. 66. The authors explain that the models consider and incorporate hierarchical structures. More description of these structures and how they are incorporated should be included in the report.

p. 66. Wording is not clear on the last sentence that begins “One of the primary objectives of these analyses was quantification of the unexplained patterns variability through...” Should it be “unexplained patterns of variability?”

p 67. “In response to this request, …through the dams in the FCRPS.” The ISAB accepts the contention that route of passage is important to understand. Therefore, either a reference to the experimental design, objectives, methodology, etc. (e.g., an ISRP review) or some description of the research is warranted.

p. 67. The idea of smolt collection delaying migration and lowering survival is counterintuitive to many readers. Presumably it is a good thing that collection avoids sending smolts through the turbines. A more detailed explanation of how collection delays migration and lowers survival would help. Putting the discussion into the context of migration and survival with other non-collection routes may help.

p 68. “During the smolt outmigration, … for small sample sizes (AICc).” Although not stated, this paragraph and subsequent model descriptions address the issue of correlation and lack of independence. Explanation of how this helps address the correlation issue should be added to the text.

p. 69. Given evidence that fish seem to tolerate the maximum levels of TDG observed (111% to 132%), why use the average TDG levels in the models. Would maximum (or even a 25% exceedance value) be a better measure of the aspect of TDG that is reducing reduce juvenile survival?

p. 70. The listed hypotheses provide a set of six predictions based on “Increasing voluntary spill.” While “voluntary” may have common understanding in dam operations, it is unclear what it precludes as non-voluntary. Ultimately, while voluntary spill is under USACE control, would “total” spill (i.e., voluntary plus non-voluntary) be more appropriate to frame the predictions? If not, why not? Perhaps a reference to the appropriate section of Chapter 7 (e.g., p. 256 to 266) explaining the different type of spill would be helpful here?

p. 71. “The CRSO-EIS alternatives…the four lower Snake River dams.” The biological benefits imply decreased WTT and FTT and increased juvenile survival and SARs, but this is not stated explicitly. The anticipated benefits of removal of the Snake River dams that are directly related to this analysis should be stated specifically.

p. 73. Figures 2.4 and 2.5. The predicted values for 2019-2021 should be on the same plot as the historical predictions. The separation makes it more difficult for the reader to discern the difference (or lack of difference) between FTT with the Flex Spill and FTT during 1998-2018.

p. 77. Figure 2.9. The confidence intervals for the estimated in-river survival for the Flex Spill cohorts are quite wide for some within-year cohorts. Presumably this is because of smaller sample sizes for each cohort, but no information on the cohort sizes is given. Please add sample size information to the figure or make a separate table with that information.

p. 77. Figure 2.9. Can the overall year in-river survival probability be estimated (with presumably) narrower confidence limits to compare to the model-based values? A simple weighted average of the three cohort values would be sufficient.

p. 81-82. For the Flex Spill experiment, juvenile FTT would be reduced by 0.4 day for steelhead and 1.1 days for Chinook salmon leading to increased survival (7% and 3% respectively). On the surface, this does not seem to be very substantial for moving SARs toward the 2% - 6% goal of the Fish and Wildlife Program. The Discussion emphasizes that the "modeling results and observations continue to demonstrate that the primary factors influencing juvenile fish travel time and juvenile survival are ordinal day, water transit time, and the number of powerhouses experienced during the migration." While that is correct, what do the results showing small potential increases indicate about the likely long-term consequences for improving juvenile survival? Also, adding a reference to Chapter 6 may be helpful here.

IV.C. Comments on Chapter 3. Effects of the In-river Environment on Juvenile Travel Time, Instantaneous Mortality Rates and Survival

This chapter is an update of the previous data for subyearling and yearling Chinook, sockeye, and steelhead in three reaches of the Columbia and Snake rivers: Rocky Reach Dam (RRE) to McNary Dam (MCN), McNary Dam to Bonneville Dam (BON), and MCN to BON. Most of the text is identical to the three previous CSS Annual Reports. Following the approach adopted in the 2021 report, the report defines the Upper Columbia reach from RRE to MCN. The ISAB recommended developing methods to adjust the values for Rock Island Dam (RIS) to maintain a longer analytical record. The CSS explored this suggestion but concluded that applying a constant correction factor across historical cohorts would not be appropriate because of the magnitude of bias varied across cohorts (CSS 2021). The ISAB accepts this rationale, but some future analyses may be improved by the longer time series so this suggestion should not be completely discarded and other methods explored such as cohort-specific covariates that may better predict bias.

Overall, the chapter is well written, and the analyses and results are explained clearly. The take-home messages in the Conclusions section are well crafted and useful for managers. In 2021, the ISAB recommended adding several major findings to the Conclusions and continuing to add major findings and highlighting new issues in future years. The CSS made those additions and included a new conclusion in 2022.

The CSS followed up on an ISAB recommendation in 2021 and further examined unusually low survival probabilities for several cohorts of subyearling Chinook in the MCN-BON reach. They found that low and variable survival probabilities were associated with cohorts that had low sample sizes. Shortened operation of the PIT-trawl below Bonneville Dam decreased the detection of migrating subyearling and reduced sample sizes. The CSS has restricted its analyses to samples larger than 2,500 individuals to avoid this potential bias. The 2022 Report adds a Conclusion that several actions are needed to estimate survival in the future for subyearling Chinook in the MCN-BON reach: 1) increased number of subyearlings that are PIT tagged upstream, 2) increased detection efficiency at McNary Dam, 3) extended PIT-trawl sampling below Bonneville dam through the end of August, or 4) a combination of these actions.

Editorial comments:

p. 93. Line spacing altered by text insertion.

IV.D. Comments on Chapter 4. Patterns in annual overall SARs

This chapter also is an update from previous annual CSS reports that the ISAB has reviewed extensively over time. Most of the text is identical to the 2019, 2020, and 2021 CSS Reports. Figure 4.39 was added to illustrate the relationship between SARs estimated by the run reconstruction method and SARs calculated based on PIT tags, otherwise all of the figures are identical to previous graphs in this chapter, updated with 2021 data.

The CSS project and its cooperators have reduced some of the known biases in their analyses, including distinguishing separate stocks in the aggregate run of adults based on genetic analysis from the Lower Granite Dam Adult Trap. Genetic analysis of smolts also separates hatchery and natural origin fish, which reduces errors of counting hatchery fish with intact adipose fins as natural origin fish. The chapter points out that increased spill and decreased precision of detection probability estimation create challenges for the run-reconstruction methods. Limited juvenile-detection facilities in the upper Columbia also contribute to overestimation of survival. The CSS Report would be improved with more discussion of alternatives going forward to address the challenges they have identified in this year’s report.

As the ISAB stated in last year's review, the take-home messages in the Conclusions section are well crafted and useful for managers, but they are almost identical to the past Conclusions and do not highlight any new results, insights, issues, or concerns. It would strengthen the Conclusions to identify any new conclusions and issues.

Also, as stated in last year's ISAB review, we suggest that more sophisticated analytical tools may be available and could strengthen the analyses. Current reports still include estimated correlation coefficients and use fairly simple statistical models, albeit with lots of intricate data manipulation. Now, with over 20 years of data, more sophisticated statistical tools might be able to characterize underlying patterns, perhaps using time series approaches. There is an apparent long-term decline, but also some evidence of a potential periodicity in these data. Once these trends are accounted for, would a pattern in the residuals become more apparent? The CSS response to our review in the 2021 Annual Report indicated that they would consider this suggestion in the development of CSS analyses for future annual reports.

Appendix B: Supporting tables for Chapters 4 – Annual Overall SARs

There have been no major changes in Appendix B. Values for 2021 have been added and overall averages or totals have been updated.

IV.E. Comments on Chapter 5. Upstream Migration Success

This chapter is updated from the previous CSS Annual reports on spring Chinook migration success with the addition of results for fall Chinook and steelhead. The contrast in inter-dam survival probabilities is small (low contrast), so the results are not too surprising. The ISAB has raised questions about the low survival from Bonneville to McNary in previous reviews. The CSS examined this more closely this year in this chapter and found that temperature in this reach was a major factor in the low survival probabilities.

The 2021 ISAB Review suggested that it would be useful to produce an overall model for Bonneville to Lower Granite dams with the same variables and compare it with the reach-specific model results to see where the largest effects occur and how these differ by species or run. The CSS does not plan to develop such a model for comparison. They feel that different processes (e.g., transportation history) determine the survival relationship in the mid-Columbia versus lower Snake. While the processes and details of the individual reaches differ, the ISAB feels that the whole-river model would provide a context for understanding the consequences of factors that influence adult survival in separate sections of the river.

The chapter indicates that "Flow, spill, spill percentage, and temperature variables were evaluated as several different metrics including daily averages, maximums and minimums on the day of arrival for each specific reach in the analysis." The 2021 report states that these variables were "variables were all taken as the daily averages." Are these variables represented differently in the model in 2022 versus 2021 or was the description in previous year's report incorrect? In several places, the text refers to "refinement of the environmental variables." The differences between environmental variables used in versions of the models in different years should be clarified.

The ISAB notes that the estimate of "age5" year effect in some tables (e.g., Table 5.3) has a very large SE, which usually indicates a failure to converge or estimates on logit() scale that are essentially "1" and are not useful (not even shown on Figure 5.4). This is why the p-value for Age 5 term is not "significant." The ISAB raised this issue in the 2021 ISAB Review, and this has not been addressed in the 2022 report. Some of the results for Age 5 are difficult to explain. Age 5 survival is 1 from Bon to MCN, very low from MCN to ICH, and then very high from ICH to LGR. Do age 5 fish successfully arrive to ICH? The text in Chapter 5 does clarify previous reports and states that "while the relative variable importance of age was near one, indicating it was in nearly all the top models (Figure 5.8), none of the age categories showed a significant effect in the model averaged results (Figure 5.5)." Further exploration and explanation of the effects of "age5" are warranted.

Previous reports stated that spill was a significant predictor of upstream survival for Snake River steelhead in all three reaches. The new analysis in this year's report indicates that spill is a significant predictor in the lower reach but has little predictive ability in the upper two reaches. This should be clearly noted and possible explanations should be added.

In the 2021 CSS Report, this chapter stated that spill appeared to affect survival of fall Chinook in the Snake River reach and summer steelhead in all reaches and pointed out that this was surprising because there is little variation in spill levels provided in late summer. Based on "refinement of the environmental variables" in 2022, temperature parameters are more important predictors and spill is less predictive. The Discussion indicates that a separate modeling effort found that water year accounted for most of the variability attributed to spill. Are that modeling effort and its full results described in the chapter? If not, they should be presented at some level so readers can better understand the basis for the statement.

The final sentence in the Discussion states that "Additional work showing temperature threshold effects on survival will be included in the final draft of this report." None of this is discussed in the chapter, so we anticipate that the methods, results, and discussion of this addition will be provided in a future draft. Will the ISAB have an opportunity to review this added material before the Final version of Annual Report is released?

There are no Conclusions for Chapter 5, and the Conclusions for this chapter from the 2021 CSS Annual Report were not carried forward, as they are for other chapters included in previous reports. Is this an error? Does the CSS plan to produce the Conclusions after completing the additional analysis of temperature threshold effects on survival that they indicate will be included in the final version of the 2022 CSS Report?

The ISAB appreciates the responses of the CSS to address previous suggestions. For example, the 2022 Report includes figures of standardized coefficient estimates with 95% confidence intervals for model-averaged effects, as the ISAB requested for this chapter in 2021. This chapter is close to becoming a "mature" chapter with relatively straightforward incremental updates in future years as new data become available.

Minor comments:

p. 229. In the next to the last paragraph, an incomplete sentence — "While our update to the environmental variables used." — was inserted. If this is simply a typo, it should be deleted. If it refers to using new or updated variables, the meaning of the addition should be clarified.

Editorial comments:

p. 191. "To improve model fit, all covariates were standardized prior to model fitting." In theory, standardizing covariates should have NO impact on model fit. But standardization often improves numerical properties of the fit (i.e., less correlation among covariates, less numerical instability of estimates, expressing effects on a common (SD) scale, etc.). Some rewording is needed here.

p. 191. "Initial model runs were conducted to determine the best fitting environmental covariates (min, max, average), of which the best performing metrics were used in subsequent modeling efforts." How was this done? Additional details and explanation are needed here. For example, was R^2 used to decide among the min/max/average?

p. 191. The model estimates the logit(S) and not S (see paragraph above the equation).

p. 193. Figure 5.1 (and others). Y-axis uses SPCH code (spring Chinook), but this acronym was not defined in report.

p. 197. Some tables need improvement. For example, in Table 5.3 an asterisk indicates "significant" estimates, but bolding is likely better.

p. 196. Replace the period at end of first sentence with a comma in the following text: "Despite this, however, a history of juvenile transport having one of the largest negative effects on probability of successfully converting remained well supported and a significant predictor of survival in all reaches. Indicating that a fundamental characteristic of transported fish is lower survival rates through the entire FCRPS, regardless of origin, independent of other factors."

p. 198. Figure 5.5 (and 5.6). Why is there no CI for the Age4 effect? (This maybe an artifact of the xlim() argument in ggplot that discards values outside of the range. Use the coord_cartesian() command to avoid this problem). Why is the order of the ages in Figure 5.6 "Age5," "Age3," "Ag4"? This likely is a typo in last label.

p. 198. Page breaks moved the legend for Figure 5.6 to the next page. Also, current document has the figure number as 4.6 rather than 5.6.

p. 199. Ditto about faulty page breaks for Figure 5.8.

p. 202. Similar comment about acronym SUCH as for previous comment about SPCH.

p. 205. Insert space after temperature in line 9.

p. 206. Figures 5.10, 5.11, and 5.12 are missing CI for Age 4 (see above comment for page 198).

p. 211. Similar comment about FACH acronym.

p. 221. Similar comment about SUST acronym.

IV.F. Comments on Chapter 6. Estimating SARS Needed to Meet Columbia Basin Partnership Goals: Three Approaches

The authors use three complementary approaches to address the relationship between NPCC benchmarks (SARs) and Columbia Basin Partnership (CBP) benchmarks (escapement) for Snake River spring/summer Chinook salmon. It was somewhat difficult to ascertain the connections among the approaches. Approach 1 compares the relationship between escapement and SARs at the basin level and suggests that basin-wide wild smolt production is about 2 million fish (the beta1 estimates on page 234). Is this a reasonable estimate for the basin? In Approach 2, smolt production is modeled at the population level (using a Beverton-Holt model). Given a smolt production, the authors calculated the SAR needed to reach the CBP benchmarks. Approach 2 also looks at smolt production at the aggregate, basin level, but it is difficult to determine the precise way this was done. The methods of aggregation should be documented, as this is needed to understand the data shown in Figure 6.4 (total yearly smolts versus total females available for natural reproduction). There may also be value in relating the results of the aggregate analysis of Approach 2 to the results of Approach 1. The objectives of Approach 3 were more difficult to ascertain, as the Beverton-Holt model used was the same as that used in

Approach 2. It seems that Approach 3 focuses on how changes in the Beverton-Holt model (modelled and empirical) affect smolt production and spawner abundance at equilibrium, and therefore the required SARs to maintain spawner abundance at this equilibrium. A schematic that clarifies the relationship between the three approaches is recommended.

Overall, Approach 1 shows that SAR and CBP benchmarks are consistent. Approaches 2 and 3 indicate that SAR targets vary by tributary, because some tributaries have higher habitat quality or quantity that support greater smolt production. Direct habitat improvements will likely help to achieve SAR targets and the CBP escapement goals, but post-smolt factors (SAR improvements) cannot be ignored. The major conclusions of Chapter 6 are that none of the CBP goals will be achieved without management to improve SARs and that CBP escapement goals for Snake River spring/summer Chinook salmon will not be achieved without management to improve SARs. Tributary-specific habitat improvements will not be enough to meet those goals.

Substantive comments:

Approach 1:

The current model for Approach 1 needs to be reconsidered for several reasons.

In general, it is hard to assess the appropriateness of the zero-intercept linear model without the data also being provided in the chapter. This is particularly true since the authors appear to be predicting for values beyond the range of the data that is used in the model (e.g., Figure 6.2).

The model assumes that the error term has constant variance. Figure 6.6 shows that this is not true. For example, an error of 1000 fish in the escapement when the mean predicted escapement is 10,000 fish is given much less weight than an error of 10,000 fish when the mean predicted escapement is 100,000 fish despite that both errors are the same relative proportion. A log(escapement) vs SAR or log(escapement) vs log(SAR) would be more appropriate. Similarly, the discrepancy measure should be based on log(escapement).

Both the Y and X variables have uncertainty. This implies that this may be an error-in-variables problem if the uncertainty is large. The uncertainty in the SAR can be quantified and included in the model if a Bayesian approach is used.

Raymond (1988) reported smolt-to-adult return for wild and hatchery steelhead and Chinook salmon from smolt years 1964 – 1984. The paper explicitly “assumed a constant rate of harvest … in Alaska and British Columbia.” He acknowledged that changes in ocean conditions might play a role “but the relative importance of each of these factors [El Niño, upwelling, changes in predators and prey] cannot be accurately determined.” It is now widely recognized that ocean conditions have changed markedly since the period reported by Raymond, and there have been changes in fisheries management as well (i.e., Pacific Salmon Treaty). Given these changes, and others, it is not clear to the ISAB that the results extracted from Raymond's (1988) as they are being used in the analysis in this chapter are appropriate. The chapter states that "These time series were used to examine the association between adult escapement and SARs prior to, during, and following complete development of the four lower Snake River dams (Ice Harbor, Lower Granite, Lower Monumental and Little Goose)." This suggests that the Raymond results are being used as a form of baseline, despite the issues associated with SARs. As the ISAB has noted before, SAR is a convenient metric for the in-river passage but not a good way to measure survival at sea because it does not explicitly separate natural mortality from fishing mortality. Explicit acknowledgement of that in the chapter is needed. Also worthy of consideration would be using CWT and PIT tag data in combination with the Raymond data.

The chapter indicates that the SAR values for the Raymond data are estimated by dividing escapement by the estimated number of wild smolts. This will induce a high correlation regardless of the underlying relationship (Jackson and Somers, 1991). This may be less of a concern for the analysis in Approach 1 because SARs were only estimated for the subset of fish that were PIT tagged. The authors should evaluate this concern. A log()-log() fit without intercept may correct this problem since the common divisor will be subsumed in the intercept.

As noted below, the role of *C (*conversion probability) needs to be clarified to make this approach comparable to the other approaches.

Approaches 2 and 3:

The report needs to explain and justify why a Beverton-Holt relationship is used instead of other models (e.g., a Ricker relationship). Further, the form of the equations for Approaches 2 and 3 appear to be the same with one important difference. For example, the equation at the bottom of p.239 is:

$$\phi = \frac{1 + \beta n}{\alpha C}$$

where $\phi$is the smolt-to-adult return rate. On page 254, the chapter indicates:

$$r = p^{-1} + Sk^{-1}$$

where *r* is the smolt-to-adult return rate, with the relationship $\alpha = p$ and $\beta = \frac{p}{k}$. If the latter are substituted into the equation on p. 239, we obtain:

$$\phi = r = \frac{1 + \beta n}{\alpha C} = \frac{1}{\alpha C} + \frac{\beta n}{\alpha C} = \frac{1}{pC} + \frac{n}{Ck} = \frac{1}{pC} + \frac{S}{Ck}$$

The only difference is the presence of *C* (the conversion probability). Consequently, the two approaches likely could be merged since they use the same fundamental Beverton-Holt formulation and methodology.

The explicit presence of *C* in the Approach 2 but not Approach 3 (and likely is hidden in Approach 1) needs to be clarified carefully in this chapter. What is being predicted? Is it the actual number of fish at equilibrium or the number of fish after adjusting for up-river survival and/or harvest? Otherwise, there may be some discrepancy between the various results depending on how up-river survival and harvest are handled and to which point the SAR is measured, e.g., at Bonneville or other upriver dams (see the third paragraph on p.240).

There are some potential issues with Figure 6.4. Assuming that it is reporting smolts per female rather than the other way around, then 1 million smolts from 10,000 females seems low. If we take 5000 as the average fecundity, we get 50 million eggs. A reasonable rate of egg to smolt survival is 0.1 based on a number of studies, and that would indicate more like 5 million smolts. The conclusion from the analyses reported in the chapter that impaired habitats need higher SARs to be viable certainly makes sense, but are all the productivities so low? Could the authors report average smolts per female or egg to smolt survival rates? Admittedly, these simple values neglect density dependence but they would still be useful to have in tables for the major populations, years, etc. Also, what proportion of the habitat would be classified as "relatively pristine" versus "poor" in this context? The authors list Copeland et al. (2021) as reporting smolts per female but this paper's full citation seems to be missing.

Including estimates like smolts per female to indicate habitat quality and productivity would be helpful, as they seem to suggest low freshwater productivity. The conclusion might benefit from an explicit statement that whole life-cycle productivity depends on egg-smolt (i.e., freshwater) productivity, survival down to the ocean, natural mortality at sea, and fishing pressure. One can argue that SARs need to improve but alternatively, freshwater productivity also needs to improve. Welch et al. (2020) indicated that Columbia Basin marine survival rates are not anomalously low for Chinook salmon. Perhaps a closer look at the egg-smolt stage is warranted to better document and clarify how and by how much freshwater productivity could be increased? An important conclusion drawn from Approach 2 results is that some populations will also require increases in freshwater productivity (i.e., habitat) as well as increases in SAR to achieve the CBP goals. This is major statement and must be better explained how it emerges from the analyses in this chapter.

Approach 3 is difficult to follow because of the presentation (equations in-line, one paragraph, etc.). How does approach 3 compare to Approach 2 given the similarity in the final estimating equations noted above?

Figure 6.7 needs further explanation. It is not clear how it comes out of the equilibrium analysis of the model. It seems the authors fix productivity and capacity that determine the Beverton-Holt function and then assume an SAR value (0.5, 1, 2, 4, and 6%) to solve for M and S (target). If any of these three parameters (productivity, capacity, SAR) are changed, a point on the plot is obtained.

More details are needed for Figure 6.9. Many readers will not be able to understand this plot.

Discussion:

The Discussion has very useful information about the three approaches, although the text does not address issues related to the data sources and some other critical information. The text in the Discussion section is a very good start for documenting the three approaches to enable readers to properly interpret the results of the three approaches. However, a more detailed and complete set of tables is needed, and these should be presented in the beginning rather than at the end. The documentation table(s) should show across the three approaches: (1) the specific questions being answered for each approach, (2) operational definitions of the state variables (stages), including their relation and estimation from the data, (3) all of the data sources side-by-side and not just citations, including if necessary an information tree diagram showing the key sources of the data used to derive the data used in the approaches, (4) overlapping and distinctive assumptions, (5) time periods used and why, and (6) prediction variables and how the results are presented. If the model developers have difficulties preparing these tables, it indicates how challenging it will be for the readers to judge the veracity of the stated conclusions derived from the results and to properly interpret the results.

The Discussion additionally should acknowledge key assumptions and caveats and provide details. For example, the analyses do not deal with parameter uncertainty, and the three models used in the approaches are general (contain a minimal number of parameters) but may also restrict how localized the results can be. Synthesis of the three approaches is needed so that common, alternative, and unique assumptions to the approaches are clear to readers. Acknowledging the caveats could help to direct future refinements of the analytical approach.

The authors do not state ultimately which of the three approaches should be given the greatest weight when evaluating the results. Perhaps, the authors chose to present the science without judgment, but it would be valuable if they were to at least offer which they consider to be strongest approach. The questions being addressed by the three approaches overlap so there is the potential to reduce uncertainty by using the multiple approaches to inform answers to some questions, albeit perhaps with unequal strength.

This chapter lacks a Conclusions section. The chapter is an important addition and it is important for the CSS to provide explicit conclusions. For example, the chapter states that

“SARs in the Snake Basin in recent years have rarely met NPCC 2%-6% goals, and spawning abundances have been predictably low as a result, but research on hydrosystem action effectiveness indicates that three-fold to four-fold SAR increases are possible with aggressive management actions through spill and breach.” This critical finding is central to several major current and ongoing management issues in the Basin.

Another important conclusion to consider explicitly stating in a Conclusions section is related to the NPCC goals of SAR values of 2%-6%. The chapter indicates that the results of the three approaches “are related to, but different from, the NPCC SAR goal of 2%-6%,” which were first suggested by PATH as a range of SARs that would allow for persistence and recovery of Snake River spring/summer Chinook salmon.

Other critical conclusions to consider for a new final section of Chapter 6 are:

- “increased SARs are required to meet CBP abundance goals for Snake River spring/summer Chinook salmon.”
- “Some populations addressed within the CBP goals will need management to increase egg-smolt productivity in order to achieve the desired number of adults.”
- “it is clear that no CBP goals will be achieved without management to improve SARs.”

Most of these are contained in the Summary section, but other chapters provide final Conclusions. A consistent structure of the chapters throughout the report would be useful for the Council and other readers.

Minor:

There are issues over the use of SARs (smolt-to-adult <u>return</u> rate) rather than smolt to adult <u>recruitment</u> rate, where recruitment is the number adult salmon that survive adulthood, with some being caught in fisheries and others returning to freshwater (or dying in-route). The former (return rates) ignores fishing mortality while the latter (recruitment rates) includes fishing. If fishing mortality varies over time, then this will not be captured by SARs, and therefore may falsely attribute benefits (or losses) to habitat changes and enhancement. Are estimates of exploitation for these fish or other similar populations available, and do they vary over time?

The chapter should discuss how jacks were handled in the three approaches and specify where jacks were included or excluded. For context, please report on the frequency of jacks for Snake River spring/summer Chinook salmon. Similarly, it was not always clear how the sexes were treated (some metrics use males and females, some just use females).

The key assumption in Approach 1 that smolt production is not limiting (second paragraph of summary) should be stated in the text. Similarly, the fact that projections are made beyond the range of the data should also be stated in the text for Approach 1.

p. 234. R^2 is not a sensible measure of fit for no-intercept models. See https://stats.stackexchange.com/questions/26176/removal-of-statistically-significant-intercept-term-increases-r2-in-linear-mo

p. 235. Please explain the sources of the data more thoroughly. McCann et al. (2022, Appendix Table B.1) is given as the source of the SAR data for the second plot but is missing from the literature cited. McCann et al. (2021) is listed and was published in 2022. However, Appendix Table B.1 provides "%SAR Estimates for 1994-2019" so it does not appear that these are the data. Figure 4.1 in this report provides a graphical summary of SARs.

The appropriateness of using steady-state equilibrium results is difficult to evaluate without seeing comparable presentations of time series plots of the data (inputs and conditions) and model outputs. Such plots, across all three approaches, should be added to the chapter.

p. 236. Table 6.2. The SARs needed to obtain the escapement goals are estimated by using an inverse-prediction (moving from Y to X as shown in Figure 6.2). It is possible to estimate an uncertainty bound by using the points where the Y-value hits the upper and lower confidence bound on the fit. We also suggest a log(Y) vs X fit as noted previously.

p. 241. More explanation would help justify using redd counts to scale the goals down to the population level. Indeed, the modeling seems to assume the redd counts are not an adequate measure to compare to the goals – if they were, these models would likely not be needed. However, redd counts are considered by the authors to be adequate to down-scale and localize the goals.

p. 242. The scenario names in Approach 2 (Tables 6.2, 6.3, and 6.5) are related to trap locations, presumably those shown on the map (Figure. 6.3). The map is useful because it not only shows the locations of these traps but also habitat "status." We suggest the traps be numbered in Figure 6.3 and these numbers (or letters or symbols) be used in the three tables rather than the unwieldy names. Also, why not order the scenarios in the tables in "ascending order of anthropogenic impact," the same way they are discussed, rather than alphabetically? Making these changes would make the tables easier to understand and more useful.

p. 246. Figure 6.4 should be explained in more detail. The axis labels should use the notation from the report; for example, it appears that the x-axis is n'(t) and is the metric for the goals and the y-axis is n''(t). To estimate the equilibrium values for a specified SAR, does one go the dotted line and read the x and y values? Or does one take a goal abundance, look it up on the x-axis and determine what SAR is needed?

p. 253. Table 6.7. The CBP goals for Snake River in Approach 3 do not match the numbers used in Approaches 1 and 2. Please reconcile.

p. 260. “until large scale reversal in habitat degradation shifts the tributary production dynamics to where it is less dependent on SARs to maintain abundances, salmon populations will continue to struggle.” And… “Significant increases have also been observed following years of favorable spawning conditions or good ocean conditions.” To evaluate the benefits of habitat improvements, why not recommend some sort of experimental approach with control and impacted tributaries? Are there data to support the second statement?

Editorial comments:

This chapter requires proofreading to ensure figures and tables are correctly cited and have accurate captions.

Some of the content could be better organized (e.g., place all introductory material in the Introduction).

Consistent notation for data, variables, and parameters should be used across the three approaches.

p. 235. Figure 6.1 is mislabeled as Figure 6.6

p. 235. Figure 6.1. The plots may show that a linear fit is not appropriate, especially for the second graph. It would be useful to identify the years for each data point to understand if there is any time series effect.

p. 236. Table 6.2. The table includes too many decimal places for the predicted SARS. Similar comments on other tables.

p. 237. R^2 for no-intercept models is not meaningful. See previous comment.

p. 237. The paper needs to be more explicit in the legend and text that the shaded regions, which indicate the variability around the fitted line (i.e., for the MEAN escapement give a particular SAR), are much more optimistic and different from prediction intervals for individual escapement values. Similarly, the explanation is needed for the inverse projections for the MEAN escapement to MEAN SAR.

p. 238. A simple flow chart would help the reader understand the four stages in the recursive model.

p. 238. Except for one sentence on page 127, Chapter 6 is the only chapter in the draft report that uses the term "natural population." The chapter should explain the terms "natural population" and "hatchery population," and it should distinguish these terms from "natural-origin fish" and "hatchery-origin fish" if their meanings differ. Do "natural populations" include natural-origin fish and hatchery–origin fish that spawn in natal streams and rivers? We assume "hatchery populations" only include fish spawned in the hatchery.

p. 239. The notation of n', n'', and n''' should be changed since it is easily confused with first, second, and third derivatives used in differential equation models. Suitable subscripts would be a better notation.

p. 239. The phrase "post smolt survivals" is vague.

p. 239. Why not use SAR in the equations rather than introducing another symbol (phi)?

p. 240 and 244. Where did the values of *C* come from?

p. 242. This is the second Table 6.2. Renumbering is needed.

p. 243. Table 6.2. What is the "spatial factor"? How are the scaled goals obtained from the number of redds (please give an equation)?

p. 243. Table 6.3. Comment in the text on why the 1950's escapements were as low as they were (i.e., near the low goal).

p. 247. Too many decimal places are reported (at most 1 decimal place should be reported) given the many assumptions and uncertainties in the model. For example, change 3.13% to 3% or ~3.1%. This needs to be considered for many of the tables in the chapter.

p. 251. Add confidence intervals for the data in Figure 6.5 so that the precision of these values is more apparent.

p. 255. Legend for Figure 6.7 indicates that upper three curves refer to the 2%/4%/6% recommended SARS, but the legend should state that they represent the equilibrium abundances of smolts and spawners at the NPCC recommended range (6%, 4%, and 2% SARs, respectively).

p. 256. Table 6.8. First column should be labeled as SAR.

p.256. Middle of page says "Catherin Creek" (missing an "e").

p. 257. Table 6.10. Too many decimal places reported (at most 1).

p. 258. Figure 6.9. Unable to tell what line corresponds to what value of *p*.

p. 259. Figure 6.10. Unable to tell which line corresponds to what value of *p*. There appears to be some extra symbols beside the word "Recent." Is this a legend that should appear outside of the plot?

IV. G. Comments on Chapter 7. Hydrosystem Changes and Their Influence on the CSS, Part I: Spill

This chapter presents a clear and quantitative summary of the history of the operations, a thoughtful interpretation of the mechanics driving those operations, and clear conclusions regarding the impacts (or lack of) of spill on salmonids in the CRSRO. The authors effectively used their data to support their interpretations at appropriate scales and show a deep and nuanced understanding of the concepts driving reservoir operations. The chapter describes the evolution of the program over time, including new methods and metrics developed for this program, and how they have contributed to environmental reviews. Ultimately, the study finds that actual spill volumes have not changed radically over time despite many changes in regulations regarding how spill limits should be defined, and thus the impact on SAR is not discernable relative to variability in river discharge. However, the study finds that high relative spill at low flows can reduce juvenile encounters of the powerhouse and have other indirect benefits to fish.

The ISAB appreciates the CSS's response to our previous recommendations to develop such a chapter. The explanation of why the structure of the chapter differs from some aspects of our recommendations is useful context, and we feel that it addresses the overall intent very well. This chapter should be a living document (e.g., updates every 5 years).

While well beyond the scope of the CSS, the ISAB suggests that similar summary chapters be written for hatchery operations; for habitat restoration; for avian predation; for seal and sea lion predation, etc. to keep a summary record of major changes to the basin in one easily accessible place. These could be commissioned work by other lead entities with expertise in those areas.

Substantive comments:

The chapter states (p. 279) that "Wild populations and hatchery groups in the Snake River basin still dominate the dataset primarily because this area was the initial focus of the CSS and funding has limited the incorporation of stocks in the upper Columbia River." Is this focus consistent with the needs of the Fish and Wildlife Program, given the listing of spring Chinook salmon in the Upper Columbia River and recent efforts by the Upper Columbia United Tribes to reintroduce Chinook and steelhead into the blocked area above Chief Joseph and Grand Coulee dams? What are the implications of the uneven focus of the CSS efforts in the river above McNary Dam? This should be addressed in the Discussion section.

Minor comments:

This chapter has in its title a "Part 1: Spill." What are the plans for Part II? A roadmap in the beginning of the planned topics in subsequent years is needed

p. 266. "Tracking spill is an important means to understand the effects of hydrosystem operations on migrating fish because it influences the probability that a fish encounters a powerhouse and is available to be transported, depending on the dam." This links to the PITPH variable (page 282), and this linkage should be noted here.

p. 284. "The most recent and most significant change in voluntary spill has occurred in the past three years and the potential influence on SAR will be unknown until adult returns are complete." The ISAB looks forward to the results of this "natural" experiment in upcoming years.

Editorial comments:

p. 268. Figure 7.2. How is the median spill computed for each period? Is this the median of the yearly medians, or a median over all values in the period?

p. 270. Figure 7.3 (and other similar figures). Add vertical or horizontal lines to represent the 3 periods similar to Figure 7.2.

p. 279. Replace "hydro management" with "hydrosystem management."

p. 279, "conversion rates of adult salmonids." Long time readers of the CSS will know that "conversion" represents survival of adult salmon as they return up-river to spawn, but other readers may not know this. It is also a conversion probability (rather than a rate, since there is no time element involved.) This should be defined when used in each chapter. A glossary would help.

p. 279. Replace "system level descriptor" with "system-level descriptor."

p. 283. The chapter states "The CSS hypothesizes that cumulative experience in the hydrosystem can have delayed effects. Hence, CSS has adopted Smolt-to-Adult Returns (SAR) as the ultimate measure of success. The CSS has performed assessments that evaluated the relationship of various survival rates to hydrosystem operational conditions while considering the influence of varying environmental conditions (Schaller et al. 2007)." First, it would be useful to provide a reference for the CSS's hypothesis about latent mortality. Second, it would be useful to explain why other common metrics are not used, such as recruits per spawner.

Would it be possible to describe the trends in juvenile survival and SARs or Chinook, sockeye, and steelhead for the period of the CSS studies? We realize there are challenges in estimating overall averages, but the chapter raises the issue of juvenile survival and SARs but provides no information about the pattern in these critical metrics through time. Perhaps the chapter could refer to representative figures in Chapters 3 and 5 to illustrate the performance of these metrics.

p. 283. "The impact of changes in spill on metrics such as juvenile survival and SAR are difficult to discern in the time series because dam passage route is only one of many processes that influence survival. Said another way, relatively small changes in spill occurred over a period with high variation in river discharge and highly variable ocean conditions." This is an extremely important finding that is often overlooked in discussions about spill. Perhaps highlighting this statement using "bold" would make sure it is noticed.

IV.H. Comments on Appendix A: Survivals (S_R), SAR by Study Category, TIR, and *D* for Snake River hatchery and wild spring/summer Chinook, steelhead, sockeye, and fall Chinook.

There have been no major changes in Appendix A. Values for 2021 have been added and overall averages or totals have been updated.

V. ISAB Appendix: Suggested Topics for Further Review 2011-2021

ISAB 2021-5, pages 4-7

1. Provide more robust introduction section that includes a summary of major findings, highlights new analyses, and describes recommendations for potential management applications of findings. Describe changes in annual report structure from year to year, including why chapters and analyses were dropped or added.

2. Describe major applications of the CSS data that have been published or reported over the last few years and briefly highlight the important findings that are based on CSS data.

3. Consider recent analyses conducted outside of the CSS to identify possible new analyses that would inform issues raised by these external analyses. Step back, decide on the core results that need to be presented, identify the major uncertainties in the results and how these could be addressed.

4. Explore analytical methods to adjust for biases for smolts captured and tagged at Rock Island to maintain a longer period of information.

5. Address the unusually high mortality rates of subyearling Chinook in the MCN-BON reach and include major recommendations in their Conclusions.

6. Form a working group to explore how newer computer technology could reduce the human cost of updating and reporting the CSS report.

ISAB 2020-2, pages 3-7

1. Expand the annual report's introductory section to highlight 1) an overall summary for the survival of Chinook salmon, steelhead, and Sockeye salmon in the Columbia River basin and how the SARs for the year compare to the long-term means, 2) new analyses included in the report, 3) major changes that may signal emerging management concerns, and 4) major recommendations for management of the hydrosystem that substantially alter or reinforce previous decisions or concerns.

2. Consider ways to address the spatial and temporal aspects of the effects of total dissolved gas (TDG) on acute and long-term survival, as we also recommended in 2019.

ISAB 2020-1, Review of the 2019 Annual Report's Chapter 2, *Life Cycle Evaluations of Fish Passage Operations Alternatives from the Columbia River System Operations Environmental Impact Statement* (CRSO-EIS), pages 5-6:

1. Perform a sensitivity analysis to investigate the impact of climate change for potential future flow regimes.
2. Compare results between different types of flow years and include demographic and other stochasticity in the models so that year-to-year variation in the output measures is more reflective of the response from different operations.
3. Incorporate the relationship of individual fish characteristics—such as body size, body mass, and condition factor, and date of ocean entry—to survival. The current literature is confusing (e.g., Faulkner et al. 2019 vs the rejoinder in Appendix G of the 2019 CSS Annual Report). Collaborate on joint analyses and use a common data set to resolve this issue.

ISAB 2019-2, pages 3-4:

1. Include information about the effects of mini-jacks on estimates of SARs and other relevant parameters.
2. Investigate implications of very low smolt-to-adult survivals (SARs) to hydrosystem operation alternatives and explore whether there is enough information to estimate how much improvements in habitat and other "controllable" aspects of the hydrosystem are needed to improve SARs.
3. Continue the work on the integrated life-cycle model looking at smolt-to-adult survival.
4. Continue to model adult salmon and steelhead upstream migration and consider adding information on individual covariates.
5. Consider ways to address the spatial and temporal aspects of the effect of TDG on survival.
6. Continue work on methods to estimate numbers of outgoing smolts at Bonneville.

ISAB 2018-4, pages 3-6:

1. Develop models for multiple populations that include combined and interactive effects.
2. Use the life-cycle models to investigate potential benefits on survival of management actions such as spill modification.
3. Expansion of ocean survival estimates to additional populations.
4. Include an analysis of mini-jacking and impact on SARs.
5. Include a more in-depth analysis of the PIT/CWT tagging experiment.
6. Improve the model for estimating abundance of juveniles at Bonneville.

ISAB 2017-2, pages 2-5:

1. Modeling flow, spill, and dam breach scenarios is very useful for policy makers. Consequently, it is important that all assumptions be clearly stated and that the results are robust to these assumptions. Work on testing assumptions was suggested.
2. Include other important processes in the life-cycle models such as compensatory responses and predator control programs
3. Elucidate reasons for shifts in the age distribution of returning spring/summer Chinook Salmon.
4. The graphical analysis of the impact of TDG could be improved using direct modeling to deal with potential confounding effects of spill, flow, TDG, and temperature.
5. The (new) modeling of adult survival upstream of Bonneville should be continued and improved to identify the limiting factors to adult returns.
6. The CSS report is a mature product and the authors are very familiar with the key assumptions made and the impact of violating the assumptions. These should be collected together in a table for each chapter to make it clearer to the readers of the report.

ISAB 2016-2, pages 5-6:

1. Use variable flow conditions to study the impact of flow/spill modifications under future climate change, and examine correlations between Pacific Decadal Oscillations (PDOs) and flows.
2. Examine impact of restricted sizes of fish tagged and describe limitations to studies related to types/sizes of fish tagged
3. Modify life-cycle model to evaluate compensatory response to predation.
4. Comparison of CSS and NOAA in-river survival estimates.
5. Examine factors leading to spring/summer Chinook Salmon declines of four and five-year olds and increases in three-year olds.

ISAB 2015-2, pages 4-5:

1. Use SAR data to examine both intra- and interspecific density dependence during the smolt out migration and early marine periods
2. Propose actions to improve SARs to pre-1970s levels
3. Explore additional potential relations between SARs and climate and ocean conditions
4. Consider ways to explore the variability of inter-cohort response

ISAB 2014-5, pages 2-3:

1. Hypotheses on mechanisms regulating smolt-to-adult return rates (SARs) [update from 2013 review]
2. Life-cycle modeling questions and Fish and Wildlife Program SAR objectives [update from 2014 review]
3. New PIT/CWT study

ISAB 2013-4, page 1:

1. Hypotheses on mechanisms regulating smolt-to-adult survivals (SARs)
2. Life-cycle modeling questions and Fish and Wildlife Program SAR objectives
3. Data gaps
4. Rationalization of CSS's Passive Integrated Transponder (PIT)-tagging
5. Publication of a synthesis and critical review of CSS results

ISAB 2012-7, pages 2-3:

1. Evaluate if the NPCC's 2-6% SAR goals and objectives are sufficient to meet salmonid species conservation, restoration, and harvest goals
2. Development of technology to improve PIT-tag recovery in the estuary
3. Review estimation methods for smolt survival below Bonneville Dam through the Columbia River estuary using PIT-tags, acoustic tags, and other methods
4. Examine measurement error in SAR estimates associated with PIT-tags

ISAB 2011-5, page 2:

1. Influence of mini-jacks on SARs
2. Effects that differential harvest could have on the interpretation of hydropower, hatchery, and habitat evaluations
3. Extent to which PIT-tag shedding and tag-induced mortality varies with species, size of fish at tagging, tagging personnel, and time after tagging

VI. References

Copeland, T., D. Blythe, W. Schoby, E. Felts, and P. Murphy. 2021. Population effect of a large-scale stream restoration effort on Chinook salmon in the Pahsimeroi River, Idaho. River Research Applications 37: 100-110. https://doi.org/10.1002/rra.3748

Columbia Basin Bulletin. 2020. Oregon Governor expresses support for lowers Snake River dam removal. February 18, 2020. https://cbbulletin.com/oregon-governor-expresses-support-for-lower-snake-dam-removal-must-mitigate-potential-harm-to-vital-sectors/

Columbia Basin Bulletin. 2021. Idaho U.S. Rep. Simpson proposing sweeping $32 billion 'Columbia Basin Fund' to finance lower Snake River dam breaching. February 21, 2021. https://cbbulletin.com/idaho-u-s-rep-simpson-proposing-sweeping-32-billion-columbia-basin-fund-to-finance-lower-snake-river-dam-breaching/

Columbia Basin Bulletin. 2022. Ibslee, Murray release "Lower Snake Dams Benefits Replacement Report," "It's clear that breach is not an option right now." August 26, 2022. https://cbbulletin.com/inslee-murray-release-lower-snake-dams-benefits-replacement-report-its-clear-that-breach-is-not-an-option-right-now/

Jackson, D.A. and K.M. Somers. 1991. The spectre of "spurious" correlations. Oecologia 86, 147–151. https://doi.org/10.1007/BF00317404

NMFS (National Marine Fisheries Service). 2020. A vision for salmon and steelhead: goals to restore thriving salmon and steelhead to the Columbia River basin. Phase 2 report of the Columbia River Partnership Task Force of the Marine Fisheries Advisory Committee. Portland, OR. https://s3.amazonaws.com/media.fisheries.noaa.gov/2020-10/MAFAC_CRB_Phase2ReportFinal_508.pdf?null

NMFS (National Marine Fisheries Service). 2022. Rebuilding interior Columbia River Basin salmon and steelhead. National Marine Fisheries Service report, September 30, 2022. 42 pages. https://media.fisheries.noaa.gov/2022-09/rebuilding-interior-columbia-basin-salmon-steelhead.pdf

Storch, A.J., H.A. Schaller, C.E. Petrosky, R.L. Vadas, B.J. Clemens, G. Sprague, N. Mercado Silva, B. Roper, M.J. Parsley, E. Bowles, R.M. Hughes, and J.A. Hesse. 2022. A review of potential conservation and fisheries benefits of breaching four dams in the Lower Snake River (Washington, USA). Water Biology and Security. 100030, ISSN 2772-7351. https://www.sciencedirect.com/science/article/pii/S2772735122000440

Welch, D.W, A.D. Porter, and E.L. Rechisky. 2020. A synthesis of the coast-wide decline in survival of West Coast Chinook Salmon (*Oncorhynchus tshawytscha*, Salmonidae). Fish and Fisheries 22: 194-211. https://doi.org/10.1111/faf.12514