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 2020 Annual Report

Members

Courtney Carothers
John Epifanio
Stanley Gregory
Dana Infante
William Jaeger
Cynthia Jones
Peter Moyle
Thomas Quinn
Kenneth Rose
Carl Schwarz, Ad Hoc
Thomas Turner
Thomas Wainwright

ISAB 2020-2
October 29, 2020

ISAB Review of the Comparative Survival Study (CSS) Draft 2020 Annual Report

Contents

ISAB Review of the Comparative Survival Study (CSS) Draft 2020 Annual Report

I. Review 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 (CSS) draft annual reports. The ISAB has reviewed these reports annually beginning ten years ago with the evaluation of the CSS's draft 2010 Annual Report and, most recently, with the draft 2019 Annual Report. For the 2019 review, there was a follow-up 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), which was not available at the time the ISAB reviewed the draft 2019 Annual CSS Report.[1] This ISAB review of the draft 2020 CSS Annual Report is the ISAB's eleventh review of CSS annual reports.

II. Summary

This ISAB review begins with an overview of the latest report's findings (this section). The review moves on to suggested topics and analyses for CSS to consider in future reports (Section III), and then lists general and editorial comments on each chapter of the draft 2020 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 2019 Annual Report, Appendix I), and we do not expect the CSS to necessarily respond immediately to new requests for further analyses.

---

[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; and review of Chapter 2 of the 2019 Annual Report (ISAB 2020-1).

This CSS Annual Report includes 25 years of smolt-to-adult return (SAR) data for wild Snake River spring/summer Chinook salmon (1994–2018). Kudos are deserved for this silver anniversary milestone. The Fish Passage Center has produced these reports since 1998, and the ISAB has reviewed them since 2010. As a result, the bulk of the report focuses on the continuation of the analysis of long-term trends. Much of the text is taken 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 identical to conclusions in previous reports.

However, many things have changed in the system over the 25 years of data collection and the impacts of these changes on the long-term analyses are largely unknown. Many of the changes in the system are summarized in Chapter 1, but the reader must infer possible impacts. The ISAB suggests a table of the changes in the system over the years, along with a brief indication of the possible impact of these changes on the estimates of the effects of the hydrosystem on salmon and steelhead survival.

The 2020 report includes a first analysis of the patterns of survival of wild steelhead in the Basin (Chapter 2). The life history of steelhead is more complex than for the other species considered. The consequences of the different life-histories of steelhead on the analysis is unclear and thus warrants a more descriptive explanation (including assumptions). The different stocks have several reaches in common in their outmigration, and a comparison of survival among the stocks in these common reaches may lead to additional hypotheses regarding possible similarities and differences. As with all initial analyses, the ISAB makes a number of suggestions on how to refine and improve the analysis.

Chapter 3 continues and expands previous years' work on the effects of the in-river environment on juvenile travel time, instantaneous mortality, and survival. An analysis to consider would be a more detailed evaluation of differences detected in these variables across individual reaches; such results could potentially suggest specific factors that may affect the populations.

Similarly, the chapter on patterns in annual overall SARs (Chapter 4) continues and extends past years' work. By now, the low level of SARs relative to the Council's 2%-6% objectives has been established. These essential but lengthy data sets and extensive summaries of results may overwhelm decision makers and the public, inadvertently giving the impression that persistently low values of SARs are inevitable. In the long-term, this can desensitize them to the potential consequences and relative effectiveness of alternative management actions that can better achieve the Council's SAR objectives. Clearly, the life cycle models of CSS and NOAA

Fisheries rely on these estimates and provide some synthesis. However, the CSS could consider developing an Impact Report, perhaps developed collectively with other groups, to communicate the most critical take-home messages for the Council, BPA, and co-managers.

Chapter 5 continues the analyses of SARs and productivity. The ISAB suggests a number of approaches to strengthen this analysis.

Finally, Chapter 6 presents a work in progress on the analysis of spring Chinook salmon upstream migration success. The ISAB makes a number of suggestions to add rigor to the analysis and to improve the reporting of results. All survival probability estimates are very high, with two segments having values of essentially 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? Are there management implications from the results of this analysis? If so, perhaps the analysis should focus on the impact of management actions. For example, might a BACI-type of analysis be informative where two reaches are compared—one as a treatment and one as an untreated reference reach (McDonald et al. 2007).

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 queries that are intended to add rigor or to enhance value of the CSS project to the Council's Fish and Wildlife Program.

Some of the past recommendations from the ISAB appear to be beyond the current scope of the CSS (see several from 2017) but will become increasingly important in the future. Some of our earlier and current recommendations may seem repetitive and unachievable within the space of 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, relevant data gaps should be identified for potential new data collection procedures, and any modifications to life-cycle models should be flexible enough to incorporate these new sources of data. This is reflected in our recommendations for future work below.

In 2020, we recommend the following topics for future reports:

1. Given the large amount of information in the CSS reports and the similarity of each year's report with previous reports, It would be helpful for the Council, BPA, Tribes,

NOAA Fisheries and other management agencies, and public to have an introductory section that highlights 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. This section could briefly identify these changes and recommendations and explain their relevance to the Fish and Wildlife Program, while directing readers to specific explanatory sections of the text.

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. Are the current data sufficient to address this problem? Are there other sources of data that would be useful? The analysis in Chapter 3 continues previous years' work and should continue. The current analysis indicates no evidence of a TDG effect (Figure 3.15). The current analysis methods must use an "average" TDG that a cohort receives and cannot identify the TDG experienced by an individual fish. There may be an issue with a lack of contrast in the TDG (see Figure 3.16) where there are only a few years with higher TDG. Could an experiment be conducted, within an adaptive management framework, at one or two dams where the TDG is varied over the course of migration? Are there other ways in which contrast can be improved?

In our review (ISAB 2020-1) 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), we recommended:

1. *Future projections of survival based on the modified flow dataset are likely to be overly optimistic. Recent years, such as 2015, have experienced very low summer flows and warm temperatures. The world's five warmest years in the 1880 to 2019 record have all occurred since 2015 with nine of the 10 warmest years occurring since 2005 (NOAA Climate.gov). A sensitivity analysis needs to be performed to investigate the impact of climate change on potential future flow regimes, such as those described in Chapter 4 on Climate Change in the CRSO-DEIS. Such a sensitivity analysis will also need to account for changes in the maturation schedule, conversion probability due to warming water, changes in the capacity-related parameters of the Beverton-Holt spawner-recruit relationship for the CSS-GR model, habitat improvements, and other factors. Do the relative rankings of the alternatives change if future climate scenarios (even if oversimplified) are represented?*

The CSS provided an extensive response to our recommendation (Appendix J). In particular:

> "Addressing climate change was beyond the scope of the task that was assigned to the CSS for this process, and the CSS agrees that climate change is important and may influence results. Because of the projected detrimental impacts of climate change, we highlighted the lower ends of the SAR intervals, and the likelihood that those lower SARs may occur more frequently in the future, in the chapter. The Action Agencies did not ask or allow the CSS to interact with them on addressing influences of climate change during the CRSO-EIS process. Appendix V of Chapter 4 of the EIS was not available to the CSS."

Similarly, the CSS indicated that their work was very much constrained by the DEIS process to focus on those elements informing the NEPA review.

2. *A more detailed comparison of results between different types of flow years would be a useful first step toward meeting ISAB recommendation (1). Demographic and other stochasticity (80 years of hydrology is a great start) should be included in the models so that year-to-year variation in the output measures is more reflective of the response from different operations.*

The CSS responses outlined how various levels of stochasticity are included in their models and have adapted their text to clarify what was modified, but no further analyses were conducted.

3. *Both models do not 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). It would be beneficial for both groups to collaborate on joint analyses and use a common data set to resolve this issue.*

The CSS provided a response where they indicated that:

> "A detailed review and re-analysis was included as Appendix G. Using the same dataset that Faulkner et al. 2019 provided, we found significant negative bias is introduced to survival estimates when using fish PIT tagged at Lower Granite Dam. We also found that Faulkner et al. 2019 ignored major drivers of bypass probability in their analysis, and their conclusion that length is one of the primary drivers of differential SARs does not explain differences in survival among populations in the Columbia Basin. ….
>
> It is important to note that a common dataset was utilized by Faulkner et al. 2019, in the development of Appendix G, and in the subsequent review and re-

> analysis submitted to Transactions of the American Fisheries Society regarding Faulkner et al. 2019. This dataset was provided online for review by the authors of Faulkner et al. 2019. To some degree, Appendix G can be considered a collaboration due to the fact that a common dataset was used. **The primary differences between the two analyses were the explicit incorporation of environmental covariates into models of bypass probabilities vs Faulkner et al. 2019 approach of modeling them as random effects** [ISAB emphasis]. Based on the data and analyses that have been conducted, we are unconvinced that fish size and condition are important factors driving poor adult returns for the upriver stocks of salmon and steelhead."

The response is helpful in clarifying the cause of the differences in the two analyses. We understand that the rebuttal will be published shortly.

In ISAB 2019-2, we recommended the following topics (*italicized*) for future reports. After each recommendation, we summarize the current status of the work to address them:

1. *Include information about the effects of mini-jacks, male Chinook Salmon that remain in freshwater and mature two years after fertilization, on estimates of SARs and other relevant parameters. Standard hatchery rearing protocols have been shown to amplify the precocious maturation of Chinook Salmon and steelhead as residuals or minijacks that do not migrate to the ocean (ISRP 2018-8, see review of Growth Modulation in Chinook Salmon Supplementation project, 200203100). Are there hydrosystem effects on mini-jack rates?*

Appendix B reports SARs with and without jacks (e.g. Table B.21), but there does not appear to be a discussion of the effects of including/excluding jacks on SARs. It must be higher (Table B.21 shows about 30% higher). What is the impact on the population trajectory of large number of jacks? In their response to our previous recommendation, the CSS plans to investigate this in their next report if funds are sufficient to allow for this activity.

Our nomenclature on mini-jacks needs clarification. Our recommendation in 2019 unfortunately blurred the definitions and used the definition of mini-jack incorrectly. Chinook mini-jacks are most commonly characterized as those fish that undergo smolt transformation, migrate to saltwater, and mature a year earlier than the typical jacks of the population, and at least two years younger than the females. Rather than "mini-jacks," the precocious male fish life history that does not migrate to the ocean is more accurately termed "mature male parr." The ISAB also uses the term "residuals" for this life history. An additional term for mature parr and residuals that do not leave freshwater is "microjacks" (Larsen et al. 2013, Hayes et al. 2015), but this causes more confusion, and we suggest only applying the term "jack" and derivatives of it to salmon that migrate to marine waters. Importantly, any of these life history variants might have very different survival rates associated with their migration or lack thereof, so it should be clear how they are handled. We recommend that the terms "mini-jack" and "mature parr" be used consistently for these two life history types going forward.

Our 2019 review of Chapter 5 on SARs and productivity recommended that the CSS should be explicit about the causes of the lack of precision in the relationship of SAR with steelhead population productivity as compared to the relationship for Chinook; we suggested that residualization could be a factor to discuss. The chapter still does not discuss residualization in steelhead and how it affects the relationship between SARs and productivity.

2. *Smolt-to-adult survivals (SARs) continue to be very low. Do we have enough information to suggest changes to hydrosystem operations that could improve SARs? Is there now enough information to estimate how much improvements in habitat and other "controllable" aspects of the hydrosystem are needed to improve SARS?*

This issue was analyzed by the CSS in their revised "Chapter 2" in the 2019 report.

3. *Continue the work on the integrated life-cycle model looking at survival from smolt back to adults (Chapter 9). A more detailed comparison with the CSS results is needed, and discrepancies need to be explained. Include individual covariates such as body mass when tagged or timing of migration to elucidate important factors that affect survival.*

The revised lifecycle models are not presented in this report because of the work for the CRSO-EIS continued into 2020 and data on newer population were not available in a form that was usable. This has delayed progress, but the CSS indicates the chapter on life-cycle models will return in 2021.

4. *Continue the work on modeling adult salmon and steelhead upstream migration and consider adding information on individual covariates such as, for example, the dates tagged as juveniles, body mass as juvenile, the dates when juveniles were detected at Bonneville, to elucidate important factors that affect upstream migration success.*

Chapter 6 in this report presents a preliminary analysis. The ISAB offers several recommendations to improve the analysis.

5. *Consider ways to address the spatial and temporal aspects of the effect of TDG on survival. Are the current data sufficient to address this problem? Are there other sources of data that would be useful? What modifications to the current data collection systems would be needed to address this issue?*

See our recommendation earlier where we suggest some additional possible analyses/experiments to improve contrast.

6. *Continue work on methods to estimate numbers of outgoing smolts at Bonneville. Could additional data be helpful (e.g., targeted releases of known number of smolts directly above Bonneville to estimate detection probabilities directly)? What are other options if the current data provide estimates with poor precision?*

This was not included in the 2020 CSS report, and the CSS indicates it is a work in progress

IV. Comments on New or Updated Analyses in the draft CSS 2020 Annual Report by Chapter

We provide substantive, minor, and/or editorial comments, as applicable, for each chapter below.

IV.1. Comments on Chapter 1. Introduction

Substantive comments

This CSS Annual Report includes 25 years of SAR data for wild Snake River spring/summer Chinook salmon (1994–2018). Kudos are deserved to the program for this silver anniversary milestone. This chapter has been reviewed many times, and we propose some refinements or additional brief descriptions to help people new to the CSS process better understand the chapter.

Many things have changed over the years (in the sampling procedures, the hydrosystem operations, and in the ecosystem), and the impact of these changes on the long-term analyses are unknown. Consider changes to the sampling procedures. How comparable are the smolts tagged earlier in the program with those later in the program? Has the PIT-tag itself changed over the 25 years, and has the size of fish tagged changed? How standardized is fish handling at the various places to ensure that handling mortality is roughly equal over time and space? There have been additions and expansions of detection arrays. Have detection efficiencies improved over time? Have there been major changes in ocean fisheries and how would this affect estimated ocean survival (BON-BON)? Ecosystem changes have occurred both in river and in the ocean. Marine mammal predation has increased over time and impacts mostly spring runs. How big of an impact on SARs has there been? Has the age structure of returns changed over time (e.g., are there proportionally fewer older fish) and is this expected to affect SAR computations? If fish are maturing at a younger age than they did in the past, an apparent increase in survival is to be expected, all other things being equal. Other examples of changes to the hydrosystem and changes to the sampling protocols are spill versus non-spill, modifications at dams and bypass structures, places and dates of tagging, tagging sites for wild and hatchery fish, and tagging at origin vs *en route*.

A complete listing of major changes (also perhaps as a timeline) would help readers track and interpret results. In addition, has the CSS conducted sensitivity analyses or tested assumptions to determine how these changes affect the estimation of SARs and other explanatory and response variables and their spatio-temporal patterns? The effects of these sampling and hydrosystem changes (which are expected over 25 years) on response variables could be contrasted to the effects of ecosystem variation (e.g., discharge, temperature at sea and in river).

A table of variables (and their units, definitions, and chapter where each is used) used in the entire report would be useful. For example, PITPH seems to be used in several ways within this annual report, as well as having different estimation methods from its use in earlier reports. Are variables like “travel time” calculated consistently throughout the report?

Development of conceptual life cycle diagrams that show how life stages use key subregions and habitat, migration patterns, and the factors that affect growth, survival, and reproduction by life stage, would be helpful in integration and synthesis and for providing a common foundation upon which readers can view the specific analyses and results. This would show the life stages and how they are considered (e.g., jacks), key processes thought to be important, whether those processes represented or accounted for in analyses, and how river operations are included. A good example is for the California Delta, which include salmon: https://www.dfg.ca.gov/erp/conceptual_models.asp

There has been a shift towards PIT tagging farther up in the watersheds to better represent dynamics at finer population scales (i.e., Major Population Group vs. ESU). This should improve the data, but such a shift also means that the fish selected to represent the aggregate population may be changing over time and hence sizes and timing might not be the same throughout the period of record. Some discussion of how this is addressed in analysis and the possible implications on long-term patterns is warranted.

How might the change in sockeye smolts from the Sawtooth Hatchery to the Springfield Hatchery (and where is that located?) be important to the interpretation of analysis results?

The description in the chapter assumes that each powerhouse experience or spill passage event are effectively equivalent at each dam. Is this assumption correct and what is the justification for the assumption? Alternatively, can the model account for any measurable differences among the experiences if these prove significant?

Editorial comments

p. 1. The term x-salt is never defined. Presumably most readers would understand this, but to be complete, please define at first usage.

p. 1. “NFH” should be spelled out at first usage.

Fig. 1.1 caption introduces abbreviations for dams. “LGR” and “BON” should be spelled out at first use in text and figure captions. (They are in the figure, but not everyone will notice that.) For the uninitiated reader, it would help to include a table describing the dams: name, abbreviation, tributary (Middle Columbia, Upper Columbia, Snake), river mile/km, and operational year. This could accompany Figure 1.1. The abbreviations are finally defined in text on p. 10.

p. 9. “Data Generated in the Comparative Survival Study.” Because various information is presented on methods, history of the program, changes to the program, etc., subheadings highlighting the types of data being discussed would help.

p. 10. Just as residualization of steelhead needs to be addressed in later chapters, so does the impact of how iteroparity can affect adult return rates. Even if residualization has small effects, it should be acknowledged and briefly discussed.

p. 13. Bootstrapping – “The method is still in development.” Can the authors be more specific about the issues under consideration? What are the issues and when (or under what circumstances) will these issues be addressed? The term “naïve bootstrap” is used referring to a specific way the bootstrap sample is constructed which may not be known to some readers – add a reference to a general definition of a naïve bootstrap.

p. 14. The two sentences beginning "Wild and hatchery smolts are marked with ...” describing the nature of PIT tags are out-of-place here; they should be earlier in the introduction where PIT tags are first mentioned.

p. 15 “Beginning in 2016, CSS tags will be provided for” Recommend substituting “tags were provided for.”

p.18, Table 1.3. Why the abbreviation “RIS” when there is space in the first column for it to be written out?

p. 19. “Fish passing through spill through the entire reach would encounter zero dams ...” Should this be worded as “would pass through zero powerhouses/bypasses”? It seems that passing over the spillway must involve “encountering” the dam.

p. 19. As we noted last year, the definition of “PITPH” here differs from later usage (“Note on the use of PITPH”), which is confusing. Here, PITPH is defined as an INDEX of the number of dams a fish passes through. In a later section, PITPH is alternatively defined as the PROBABILITY of passing through the powerhouse at a single dam. Some distinction in the two uses should be included in the acronym, perhaps use “cumulative” or “total” PITPH for the across-dam total, or add subscripts to the acronym to clarify (e.g., $\text{PITPH}_{\text{Tot}}$ for the total; PITPH_{i} for a single dam; $\text{PITPH}_{\text{ratio}}$ as the mean of PITPH_{i} across multiple dams). Also, it should be clarified that the index is the sum of the individual probabilities.

p. 19. “The CSS had historically use average spills…” Recommend substituting “historically used average spills.”

IV.2. Comments on Chapter 2 Patterns of Survival of Wild Steelhead in the Columbia River Basin

We appreciate that this chapter is the first exploratory analysis of these data; we provide some suggestions for improvement.

Substantive comments

The life history of steelhead is complex. How does this affect the analysis and conclusions? If parr residualize permanently (i.e., live out their lives as resident rainbow trout), then this confounds estimates of in-river survival. Moreover, if they delay outmigration for a year and then leave, how are such records handled, and does this occur often? Will they be included in the number of adults returning for a given year? Perhaps the proportion of smolts that have these alternate life histories is so small that this is moot, but some discussion is needed. Similar issues regarding repeat spawners (kelts) should also be addressed.

Estimates of SARs are from the first dam encountered to Bonneville and therefore survival between the point of release and the first dam is not included. Why was this first segment of mortality not included? Furthermore, the number of smolts that survive to the first dam is an estimate with uncertainty, but this uncertainty does not appear to be included in the overall uncertainty of the survival or SARs. Please explain how survival from release to the first dam and the uncertainty of the number of tagged smolts alive at the first dam are considered.

Similarly, the variable PITPH is also estimated but used as a “known” variable in the regression. If the uncertainty in the estimated values of PITPH is large, then the error-in-variables problem can become important and the effect could be a dampening of results (i.e., slopes pulled towards zero). How uncertain is the PITPH index?

The annual report should compare estimates of survival over the four dams that are experienced by all populations. While this does not portray a complete story for the upper reaches, it could provide a common reference. This will allow a consistent comparison of survival across the populations in the lower reaches.

SARs and ocean survival represent an “integration” over several years of ocean conditions because some fish return at age x, some at age x+1, some at age x+2. However, the modeling only considers ocean conditions in the year the brood-year first enters the ocean. Ocean conditions in the second year after salt-water entry could be important. Models should be extended to include the possibility of important effects from multiple years of ocean conditions. Also, do all populations have the same age composition, and is it assumed fixed? If the age compositions vary, then it seems likely that such variation will also affect raw return rates. For example, smolt size typically affects age at return.

The reporting of the results of analyses rely on significance results and reporting of the relative importance of different (standardized) explanatory variables. Interpretation should be expanded to look at the effect sizes of the explanatory variables. Knowing significance, sign of effect, and relative importance is a good initial step but leaves valuable information untapped, especially for informing management. How much does SAR or survival change in response to changes in native units of explanatory variables? Variable X1 can be the most important variable but only have small actual effect on survival. This overreliance on “p-values” is an important statistical issue (e.g., Smith 2020), and these types of analyses in the CSS could be leveraged, with only moderate effort, to provide valuable additional information.

Minor comments

For the general reader, a brief paragraph introducing the unique aspects of steelhead life history would be useful, including its relationship with rainbow/redband trout (residualism), frequency of repeat spawning, variable age structure, run timing, and ocean migration. This would provide essential background for understanding the details of the subsequent analysis.

Moreover, the opening statement says that Stock, WTT, and PH were important explanatory variables for freshwater survival. Given that these were key inputs into the model, was this unexpected? The same question applies to the ocean survival model.

The reported analysis examines the effects of ocean variables on both SARs and ocean survival. But ocean survival is a component of SAR. Were the results substantially different? Why? And what do differences in results mean?

What level of catch in fisheries or predation (e.g., sea lions) on returning adults occurs in the Columbia River below Bonneville Dam? Have these levels of removals been relatively constant over time? How will variation (presumably predation has increased over time) be expected to affect the results?

How much of the water transit time differences just reflect different distances (location A to Location B) traveled versus other factors that affect hydrodynamics? If the Snake River fish have farther to travel than John Day fish (and this seems more than likely), then is the longer water transit time for Snake River fish simply a reflection of this distance?

The explanatory variables are each standardized (X-mean/SD). This is similar to a unit change (e.g. cm to inches) and so has no “impact” on the analysis. However, the estimated coefficients associated with a standardized variable are difficult to interpret. The results with standardized variables should also be “back-transformed,” so results can be expressed in native units of the explanatory variables.

SARs are computed using detections of adults at Bonneville. We know from past CSS Annual reports that detection probabilities are quite high for adults at Bonneville. A note should be added to the report explaining this detection probability and that the number of returning adults not detected at Bonneville is likely to be quite small.

The Methow-Entiat data are pooled as a single unit even though the Methow group also passes Wells Dam. If the steelhead bypass the powerhouse and turbines, then this may not have a large effect; however, if these fish experience an additional PH, then this could be important to the analyses. Moreover, as the JDA population does not experience the MCN PH, how is this accounted for in the analyses? This should be explicitly (albeit breifly) described. This is critical given that one of the conclusions states "the number of powerhouse passage events were negatively associated with freshwater survival of steelhead groups.”

Table 2.2 shows the mean freshwater, ocean, and SARs, but the unbalanced data (i.e. not all groups were followed in all years) implies that the difference could be due to the imbalance. We suggest fitting a randomized block model that accounts for this imbalance to make interpretation easier. Also, adding measures of uncertainty to the entries in the table will make comparisons among groups easier.

Figure 2.7 shows that the predicted Freshwater Survival is relatively flat. So, the “variable importances” may be trivial as the net impact of all variables is very small. This needs discussion.

Figure 2.8 shows that Ocean Survival for the Yakima group for 2008-2010 was not predicted well. Any insights as to why? The same applies to the SARs.

Figures (e.g., 2.2 and 2.3) showing explanatory variables are shown in native units. Please add a secondary Y-axis on the right showing the standardized scale.

Editorial comments

Is it possible to show a table with each dam and associated reservoir as a row from interior to the coast, and each stock as a column, with the relevant estimate of survival for each stock and dam (averaged) shown? That would be very informative.

p. 24. Are separate linear regression models being used for each? The report later says that a generalized equation was used. This needs to be spelled out more clearly. The calculation of the explanatory variables is included (as well in other chapters), suggesting the same variable name and label may differ across chapters. For example, is water-travel-time the same everywhere? Is the general calculation the same but details are different for different analyses because WTT is tailored to each situation?

p. 24. The words “the flow data” should be deleted, or it should be changed so “the flow data were standardized by …”

p. 24. Mention is made of the winter biomass data, and it might be good to indicate that these data are from the winter prior to fish ocean entry year rather than the winter afterwards, if we understand the Daly et al. (2013) paper correctly.

p. 24. Steelhead have much different ocean migration patterns than Chinook and coho, so the Daly et al. (2013) reference (which focuses on Chinook and coho) seems to have little relevance to steelhead.

p. 26. The statement “The WTT and PH indices were standardized by subtracting the mean and dividing the SD calculated across years and stocks” is not clear. Presumably, the mean and SD of WTT and other indices would be computed across years but not across stocks.

p. 29. Capitalize "Relative variable importance.”

Figure 2.3. Can error bars be added to PITPH estimates to provide an idea of the uncertainty in these estimates?

Figure 2.4. It would be helpful to make the error bars reflect the 95% CI for easier comparisons among groups. Also, the overall means need to be adjusted for the imbalance in the release years among groups. As much as possible, make the Y axes consistent across plots. For example, the first graph is on the probability scale while the second graph show percentages.

Figure 2.5. The comments for Figure 2.4 also apply to Figure 2.5.

In many places in the chapter the term “survival rates” is used. However, the reported metric of survival is as a probability or fraction and not an instantaneous rate (not per unit time). Clarify all uses of “survival” and “survival rate” in the text.

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

As stated in the previous ISAB review of the 2019 Annual Report, the chapter is well written, and the analyses and results are explained logically and clearly. The take-home messages are well crafted and clear, and useful for managers. The 2020 report includes new data and analyses for subyearling Chinook in the Rock Island Dam to McNary Dam and McNary Dam to Bonneville Dam reaches, and new results for sockeye in the McNary Dam to Bonneville Dam reach. The chapter is an update of the previous data for 10 of the 13 species-reach combinations and most of the text is identical to the 2019 CSS Annual Report.

Substantive comments

The previous ISAB review (ISAB 2019-2) recommended that the CSS expand their discussion of the TDG results to include factors at other broader spatial or temporal scales that might obscure the local effects of TDG. For example, there may be compensatory responses beyond the reach scale or responses in one life stage may influence other life stages. The 2019 CSS Annual Report was revised to include a paragraph that explained the timeframe of the biological effects of TDG and indicated that their effects represented responses at the reach scale. The Discussion still does not discuss effects beyond the study reaches and how effects can influence later life stages of salmon and steelhead.

This chapter provided a detailed evaluation of findings detected for multiple species in multiple reaches related to variation in travel time, mortality, and survival. This may be beyond the scope of the current analysis, but it may be relatively easy to leverage more insights from the data by performing a more detailed evaluation of differences detected across individual reaches; such results can potentially suggest other factors that may be affecting these variables.

Minor comments

The Results section noted that several cohorts of subyearling Chinook “showed unusually high mortality rates during 2014-2019.” The potential causes and management implications of these unusually high mortality rates (0.4/d) are not addressed in the Discussion. The addition of subyearling Chinook salmon strengthens this chapter and warrants further attention.

It appears that Fig. 3.5 in the 2019 CSS Report is the same as Fig. 3.6 in the 2020 CSS Report. The values and error bars for H+W STH/RIS-MCN, H+W CH1/RIS-MCN, and H SOX/RIS-MCN appear to be different between these figures. The same is true for H+W STH/MCN-BON and H+W CH1/MCN-BON in Fig 3.5 (2019) and Fig. 3.7 (2020). The values for in-river survival probability for H+W STH/MCN-BON in Fig 3.7 (2019) and Fig. 3.10 (2020) differ greatly beyond different graphical scales. The scales have changed slightly, which could account for the appearance of a difference. But several of the values are clearly different with closer inspection. What are the reasons for the differences?

The effects of each PH experience is treated as essentially equal and therefore additive. Is this assumption justified? Similarly, average (mean) proportion of spill is treated such that daytime or evening and large or continuous spill do not influence the metric. Again, is this assumption justified? A brief discussion should be added.

In its review of the 2019 CSS Annual report, the ISAB noted that the graphs in Figures 3.2 to 3.7 were useful for illustrating the trends across years, but the scale and size of symbols were too small to allow readers to see within-year trends. The size of the graphs was decreased in the 2020 CSS Report, exacerbating this limitation. ISAB could not find these data in tables or in an appendix, though Appendix A did provide instructions for accessing the data through the FPC Web site. The text could refer the reader to this guidance in Appendix A.

The report clearly shows a low survival of sockeye released from the Springfield Hatchery in 2015 to 2017. This appears to be a problem with hatchery water chemistry and acclimation to river water once released. A note should be added to discussion. See this article from 2017: https://idfg.idaho.gov/press/biologists-think-theyve-found-answers-low-survival-sockeye-salmon.

Sockeye salmon condition from Springfield Hatchery 2015-17 were not included. As a future effort, would it be possible to include “fish condition” as a variable? This might be K (condition) at mark and recapture, disease status, or other biologically meaningful metric of individual fish health. Would the model for computing FTT and Z be improved by including an explanatory variable indicative of a fish encountering a predator (avian, native or non-native fish, or mammal)? Similarly, could water temperature be used as a covariate or surrogate to represent disease expression?

The authors propose other spill treatments to improve survival and return. How feasible are these? What are major impediments to these proposals? The ISAB is not looking for a long explanation, but a few lines to indicate that these are sensible experiments aimed at addressing critical uncertainties.

When discussing how to improve precision, two choices were increasing releases or improving the detection probability. These alternatives have different costs. To help decide on these options, how many more fish must be released or how much higher does the detection need to be to reduce the standard errors by, for example, a desired amount or percent?

Editorial comments

Figures - The scale for FTT and Z on the multiple graphs is not consistent. Part of the signal will be the relative FTT among reaches for species groups.

Many places. Please make the font and paragraph spacing uniform throughout this chapter.

p.39, Table 3.1 and associated text. It is not obvious why there would be fewer cohorts (FTT and Survival) downriver between MCN-BON than for either upriver reach (LGR-MCN and RIS-MCN). For example, steelhead have 42 cohorts down river, while the upriver cohorts add to nearly 180. Is this a sampling issue or other? Some text is warranted to explain.

p. 59. The term “Julian day” is still used in some sections of the text and “ordinal day” in other sections. “Ordinal day” or “day of the year” are correct in most uses – refer to Wilimovsky (1990). Please make use of these terms consistent throughout.

p. 44. In the ISAB review of the 2019 CCS report, we suggested that the data for TDG and the estimated instantaneous mortality rates could be included as tables either in the text or in an appendix. The 2020 CSS Report still does not provide this information or provide directions to a source for these data.

p. 59. The phrase “However, a number of observations … maximum of 136%” is vague. More precise description of the number or proportion of observations is warranted.

p. 63, Figure 3.15. What are the units on the x-axis for the TDG effect?

p. 64. Capitalize “relative variable importance.”

IV.4. Comments on Chapter 4. Patterns in Annual Overall SARs

Chapter 4 is a summary of annual (by smolt migration year) estimates of overall smolt-to-adult return rates (SARs) for wild and hatchery salmon and steelhead populations. Annual overall SARs are based on cohorts of PIT-tagged fish that experienced the same conditions as untagged smolts under a given year's fish passage management scenario. Populations analyzed are from the Snake River, Mid-Columbia, and Upper Columbia regions.

The analysis of annual overall SARs in Chapter 4 is mostly consistent with the 2019 CSS Report. A major addition to the chapter is the inclusion of estimates of ocean survival rates for select groups of Upper Columbia wild Chinook and steelhead, which the ISAB requested in 2019. Reach-level overall SAR values for Upper Columbia PIT-tagged groups were estimated beginning with Rocky Reach as the upstream most site.

Substantive comments

The ISAB continues to be concerned that the overall SARs of Snake River wild spring/summer Chinook and steelhead are still well below the Council's objective of 2%-6%. This is illustrated in Appendix B (Table B.1) with the 2018 migration year for LGR-to-GRA SAR: 0.33% without jacks, 0.39% with jacks and LGR-to-BOA SAR: 0.45% without jacks, 0.52% with jacks. Chapter 4 of the CSS Annual report continues to provide these estimates and point out the continued low SARs for both Snake River wild spring/summer Chinook and wild steelhead and Upper Columbia wild spring Chinook and steelhead. The CSS Reports identify incremental improvement in tagging, detection, and analyses to improve our understanding of SARs. The report should continue to strongly advise about complacency about the low SAR values.

The ISAB appreciates the effort of the CSS to develop time series estimates of first-year estuary and ocean survival rates for upper Columbia wild spring Chinook and steelhead (Appendix B, Tables B.128-130). The 11-year geometric means of the estimated ocean survival rates for wild steelhead from the Entiat and Methow rivers for 2008-2017 were greater than the estimated ocean survival rates for 1) wild spring Chinook from the Wenatchee River during 2007–2018, and 2) wild spring Chinook from the Entiat and Methow rivers during 2008–2018. These rates and relative differences between steelhead and Chinook for the upper Columbia are similar to those for the Snake River. These estimates are reported in the Results section and provided in Appendix B, but the results and their management implications are not addressed in the Discussion. The Report should expand the discussion to further explore the ocean survival of Upper Columbia wild spring Chinook and steelhead.

Editorial comments

p. 77, “Ocean survival rates….” These are a probability and not an instantaneous rate (i.e. not per unit time). Remove the word “rate” here and elsewhere. For example, Figure 4.28 uses “ocean survival rates,” but this is not a yearly value but a total ocean survival over multiple years.

Table 4.1. Some readers may not understand the notation used in the *R* statistics software, i.e., what is the difference between a model term with a colon (:) and an asterisk (*)?

Table 4.2. *Species* is an indicator variable, so the species effect is the “difference” in responses between the two species. This should be indicated in the table, i.e., this estimate is the effect of steelhead vs yearling Chinook salmon.

Figure 4.12. Check the legend. The dark line for “predicted ST” seems odd?

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

No comments from the ISAB.

IV.5. Comments on Chapter 5. SARs and Productivity

This chapter analyzes the relation between smolt-to-adult survival (SAR) and realized population productivity of wild Snake River steelhead. The 2020 CSS Report includes a major addition, the analysis of spring/summer Chinook based on data from the Snake River basin submitted to NOAA Fisheries for the 2020 ESA Status Review. The 2020 Report also includes a new data series from East Fork Potlatch River, which is intended to illustrate inter-population variation in the relationship between SARs and productivity relate in steelhead, expanding the data sets for steelhead to seven sites. The age-class composition in the East Fork Potlatch River site differed substantially from the previous six reaches, expanding the scope and representativeness of the study.

Overall, the analysis supported previous analyses and a recent paper by Petrosky et al. (2020) that indicate that 2%-6% SARs are required to maintain or increase steelhead populations.

In our review of the 2019 Annual Report, the ISAB encouraged the CSS to evaluate the potential influences of age- or size-related differences in estuary and ocean mortality on the population productivity among populations. The CSS includes estimates of ocean survival of Snake River steelhead in its reports; however, it is likely that the number of observations would be too low for analysis for several, if not all, of the populations. The Discussion section in the 2020 CSS Report expanded on this topic of the variation in the relationship between productivity and SARs, noting that the low current survival rates and the tendency to aggregate populations in migration studies limit their ability to analyze factors, such as differences in ocean survival for the different populations. The ISAB appreciates the additional attention to these relationships and encourages the CSS to continue pursuing alternatives and exploring the effects of these factors.

Substantive comments

The 2019 CSS Annual Report developed a new conclusion that the "relationship of SAR with steelhead population productivity is less precise than for Chinook and varies among populations." The ISAB recommended the CSS discuss the causes for the lack of precision (e.g., sample size, residualization). The conclusion was expanded to include the suggestion that "lower precision is likely associated with the more complicated life history of steelhead and variability among populations, which buffers some populations better than others."

A paragraph in the Discussion was expanded for the 2020 report to include potential effects of the differences in mean age compositions for different populations. The seven sites in the 2020 analysis differ substantially in age class composition. Could these data be examined to evaluate the effect of age class composition on the variation in productivity responses? We predict that varying age distributions (if stable through time) might influence productivity.

As indicated above, the text indicates that data were lacking because of low survival rates and aggregation of different populations in migration studies. Given that limitation, could the CSS suggest new studies or modification of current monitoring and data management to allow further analyses? What sample sizes would be necessary to allow the CSS to evaluate population-species survival rates for Snake River steelhead in the ocean?

The Discussion ends with the statement that "continued curation of these data will allow for more detailed analyses of spatial correlation in trends and potential effects of interpopulation variation in life history characteristics that may drive observed patterns." These data clearly are essential, but will it require more than simply continuing the curation and QA/QC of the data to allow for more detailed analyses? What changes in tagging, monitoring, and modeling would provide the most valuable advances in understanding the relationship between SARs and productivity? Addition of this information to the Discussion will provide context and ideas for additional analyses.

The analysis of spring/summer Chinook in the Snake River basin indicates that "observed productivity was generally positive (median ln Rsg/S > 0) when SARs were in the 1%–2% range."

This differs from the relationship for Snake River steelhead, in which median productivity was negative (median ln Rsg/S < 0) when SARs were in the 1% to 2% range. The text suggests that this high productivity at relatively low SARs might be related to higher productivity at low spawner densities. What were the fractions of minimum abundance threshold (MAT) for steelhead on their natal spawning grounds? This difference in the relationship between SARs and productivity can have major management implications and should be explored further and explained clearly for decision-makers and co-managers. The Report also should provide one or more statements in the Conclusions related to the greater productivity of wild Chinook populations at lower SARs than for steelhead.

For spring/summer Chinook, the relationship between population productivity and SARs was lowest for the Grande Ronde/Imnaha (GRIM) populations. This population had a significantly greater percentage of hatchery-origin spawners (11-58%) than the other three populations (mostly < 3%, with the exception of the E. Fork Salmon River at 25%). The chapter does not discuss this difference or provide additional information related to possible hatchery influences on the relationship. If this is a valid difference, then the potential effect of hatcheries on the relationship between SARs and productivity should be discussed and included in the Conclusions.

While Figures 5.1 and 5.3 are presented as part of the strictly graphical analysis, “trend” lines are fitted and presented without any supporting statistics. To the extent estimation of these lines is biased, there will be bias in the visual presentation of the results. Because standard linear regression is applied to log-transformed data that has non-Gaussian estimation errors in both the dependent and independent variables (“errors-in-variables"), bias in the fit may exist if the uncertainty in the X-values is large. How do the results differ if an error-in-variables approach is taken? If the error in X is small relative to the contrast in X (the range of X values), this concern may be moot, but a quick analysis and associated commentary would confirm this.

Age structure is computed differently for productivity (mix of annual and averaged age structure) than for SAR weightings (averaged over years). How does this affect the results? The Report states, "to calculate SARs by brood year we weighted multiple years of SARs by an average juvenile outmigrant age structure (Table 5.1)." How many years of data were averaged, and were they representative of the entire timespan of the analysis? How much year-to-year variation was there? Were any adjustments made for repeat spawners? These issues are at least mentioned in the Discussion as contributing to the errors in the analysis; further discussion is warranted.

The 2020 CSS report indicates “We expected variation for steelhead populations in their productivity response to SARs due to their complex life history. The variation is likely caused by changes in survival and productivity during periods when the fish are residing upstream of Lower Granite Dam" (p. 124). Causes also likely include errors in estimating age structure for both SARs and productivity. A more thorough discussion of the factors contributing to the variation is needed.

Regarding the statement "For years when the weir was not operated … was operable" (p. 124), the ISAB recognizes complete data sets are often difficult to achieve; however, some justification is warranted for simply assuming the average (mean) hatchery contribution. How robust is this assumption given that there is likely a high variance around these mean estimates?

Editorial comments

A simple graphic of the life history of steelhead would be useful for readers of this chapter.

p. 122. Some brief text is warranted to clarify whether the 2-6% range and 4% average (mean) for SARs is over time (years) for each local stock/run OR over stocks/runs in the given year. In short, what data are included in the individual SAR calculations should be transparent to the reader.

p. 122. Regarding statement "Additional SAR objectives … of wild salmon and steelhead," future reports might benefit with the inclusion of a table or chart that lays out what goals and objectives are linked to SARs in each of the higher-order plans they discuss, particularly for any that diverge from the Council's 2-6% goal. Perhaps these are well-known throughout the Columbia River Basin, but the statements are rather lean on details.

p. 124 (Steelhead) and p. 125 (Chinook salmon). Methods regarding the estimation of productivity (R_{sg}/S) are incomplete. Presumably, R_{sg} is the adult returns resulting from spawning in the previous generation (S), but this is never stated. Please define R_{sg} and S, and clarify whether R_{sg} is natural-origin only or includes hatchery strays. Also, please explain how age-structure is incorporated into the calculation (an equation would be helpful).

p. 128, Figure 5.1. The figure shows that approximately half of the data points fall below the 2% minimum for SARs. Moreover, only one data point exceeds the mean criterion of 4%. Also, approximately half of the data points (not the same ones as above) fall below the lnRsg = 0.0 threshold for productivity replacement. This does not appear to have an explanation as to what these mean for population growth and recovery. A brief expansion on its relevance is warranted, for example, by adding a bullet to the Conclusions to convey this along the lines of "Current conditions do not produce SARs > 2% for nearly half …"

p. 133. The phrase “Demographic data for steelhead populations … difficult to measure” needs clarification. It is unclear what the difficulties are in measuring adult escapement. Is it a problem with sampling efficacy, marking and detection, or innate biology (e.g., retention in freshwater for up to 5 years, or an uneven pattern to smolt age at migration)? Also, how will longer spawner abundance time series add “sensitivity” to the analysis? This is not obvious. Is the report referring to the power of the analysis?

p. 134. Regarding the opening statement “There are some obvious differences … exhibit a positive correlation” – the CSS report does not include analysis or reference data on r or R^2. If correlations have been calculated to support relationships, these should be included. Later in the paragraph, the report suggests that more populations will be added to the analysis to help “understand” the response of populations to variation in SARs. Suggest substituting “explain, depict, or illustrate” for “understand.”

p. 135. “Declines in life-cycle productivity of wild Snake River steelhead and spring/summer Chinook Salmon populations were associated with brood year SARs less than 1%, and increased life-cycle productivity occurred when brood year SARs exceeded 2%.” The terms "declines" and "increased" imply time trends in productivity, which is not something the analysis looked for. These terms should be changed to something like "lower" productivity and "higher" productivity.

p. 135. The text says: “... which buffers some populations better than others.” We are not sure what is meant here by "buffers" - buffers from what forces? Population buffering was not mentioned anywhere else in the report.

IV.6. Comments on Chapter 6. Spring Chinook Upstream Migration Success

We appreciate this is a work in progress and offer suggestions for improvements for the future.

Substantive comments

It does not seem necessary for the analysis to create the “pseudo-detection occasion” and some simplification in the analysis methods may be possible. The estimate of detection at BON is simply the number of fish detected at BON / number detected at BON or higher, with a standard error found using a binomial distribution. Then one can create capture histories for ALL fish detected at BON or higher as being released at BON in the first detection occasion. This approach will give all the estimates of reach survival needed for the models. Please explain or eliminate the text about the “pseudo-detection occasion.”

All survival probabilities are very high, with two segments essentially having values of 1.0. Why even perform the analysis when survival fraction (upstream spawning success) does not vary from a value of one unless the analyses are sufficiently sensitive to respond to such minor differences? Even for Bonneville-to-McNary, where survival varied the most, the range of values was only 0.729 to 0.822 for 2008 to 2019. Give the lack of contrast in many reaches (e.g. Figure 6.2), the value in the modelling exercise would seem to be very limited. This is noted in the discussion but needs more details.

The model described on page 138 needs to be more carefully defined. In the introduction, it was indicated that multiple logistic regression would be used. Then on page 137, the report indicates that a linear mixed model is used. It would appear that a generalized linear mixed model is being used for each reach. Is this correct? If so, the model should be specified as:

$$Y_i \sim Binomial(p_i)$$

$$\mu_i = logit(p_i)$$

$$\mu_i = \cdots$$

where

- Y_i is a 0/1 variable indicating if the fish survived passage in this reach
- p_i is the probability that the fish successfully survives this reach.

What data are used? For example, consider the MCN to ICH reach. Were only fish detected at MCN used or were fish detected at MCN or higher (because if detected upriver, you know the fish was alive at MCN) used? What if a fish was not detected at ICH but was detected upstream of ICH? What is done after the last detection of a fish where it either died or was not detected further? Logistic regression assumes that the fate of each fish is known with certainty. But a fish released at MCN and not detected further does not imply that the fish is dead – perhaps it was just not detected. If the detection probabilities are very high this may be moot, but no information is provided on the estimated detection probabilities from the capture-recapture model first fit to the data. And what about fish detected at BON and never seen again. It is not known if the fish was alive at MCH – was it included in the analysis for this reach?

Temperature effects were modelled as linear or quadratic effects. But there may be threshold effects (i.e., no effect below temperature *x* and catastrophic effects above temperature *y*). Analyses with a categorical temperature effects (e.g., with a breakpoint at the pejus temperature) should be investigated.

Dealing with multi-collinearity (p.138) could likely be handled using AIC directly. This should be explored.

What are the potential management implications of this analysis? Do the results imply a change in management actions anywhere? The results should be synthesized with some discussion of their broader implications for management.

Minor comments

Arrival timing was defined as the time of first detection in a downstream reach. What if a fish was not detected? Is it ignored? Similarly, the same question for travel time.

Standardizing covariates does not improve model fit (R^2). Standardization is like a change in units (e.g. m to km) so does not affect the overall fit of the model. All standardizing achieves is to make the numerical methods more robust in the fitting. Were the 0/1 covariates such as transported and wild/hatchery also standardized?

Age is included as a continuous covariate. A categorical covariate may be more appropriate because survival is likely not linear with age.

Under future plans, there is no mention of expanding the analysis to steelhead. Has this intention (response-to-comments on last year's Ch. 7) been dropped?

Editorial comments

p. 137. For clarity, replace “tagged” with “PIT-tagged” so the reader knows immediately what data set is used.

p. 137 (and throughout Results). The term "conversion" is undefined. In the Results, this term seems to correspond to “survival” in the figure and table captions. If this is the same as survival, use "survival” throughout; otherwise, please define the term.

p. 137. Please explain why squared temperature was included, while all the other variables are linear.

p. 138. "pseudo-detection occasion ... release occasion" - This jargon needs to be clearly defined.

p. 138 “… accurate survival estimates for the BON-MCN reach.” The pseudo-detection occasion gives you a detection probability at BON and the survival probability in the BON-MC reach but has no impact on “accuracy” of the estimates.

p. 139. The term "model averaged object" should be “model averaged effect.”

p. 138. The predictor variables in the equation are abbreviations that can be guessed from rereading the "Data" section, but it would be good to include a list or table defining all the variables.

p. 139. Figure and table numbers in the text do not correspond to those in the captions. For example, the reference to "Figure 1" should likely be Figure 6.1. Please correct figures numbers in the text throughout this chapter.

p. 141. Both the figures and Table 6.3 show some form of interval about the estimate, but these are not defined in the text nor in the captions. Please describe what they are (confidence intervals on the mean, prediction intervals) and how they were calculated. Some of them appear to be symmetric (which is surprising for survival estimates), while others do not. Fewer decimal places should be reported here (at most two).

p. 142. "… more likely to convert ..." How much more likely? Give the estimated odds ratio and measures of uncertainty. Ditto for all statements about effects – give estimates of effect sizes and their uncertainties.

p. 143, Table 6-4. P-values and asterisks and "significant predictors" should be dropped, as they are inconsistent with the AIC model-selection framework. See comment about page 141. In addition, too many decimal places are displayed and define 0/1 variables (i.e., does 1= wild or hatchery fish). Please align the rows so that coefficients for common effect across the reaches are more easily compared and add estimates of uncertainty. Some models have the square term but not the linear term (e.g., Temp^2 term but no temperature term); these non-hierarchical models should not be fit because they make a specific assumption about the shape of the curve (i.e., symmetric about 0).

p. 144. Figure 6.4-6.6. Please sort the variables in the same order across plots to make it easier to compare the results for the different reaches.

p. 145. The phrase "strong model support and significant correlations between a history of juvenile transport and lower survival probabilities as adults migrate upstream." Transport is a categorical nominal scaled variable and so "correlation" is not defined.

p. 145. The phrase "significantly higher probabilities" can be much more informative. Please provide estimates of effect sizes and uncertainty.

p. 145. The text "as it may be driven by a relatively small number of individuals migrating during a period of elevated temperatures" is important and should be emphasized more. The small number of individual migrating would not cause a “temperature effect” to be detected unless they all died. This is potentially an important finding.

p. 146. Check the results for harvest. This is quite unusual for a variable to appear in the top models and then have an effect size the includes the value of 0.

IV.7. Comments on Appendix A: Survivals (SR), SAR by Study Category, TIR, and D for Snake River Hatchery and Wild Spring/Summer Chinook Salmon, Steelhead, Sockeye, and Fall Chinook

This is an update of the appendix from previous years.

The presentation of the results graphically suggests it might be interesting to look for cross-stock patterns of coherence and dissonance. Are good years good for several stocks (synchronous)? This is done in the text, for example, p.99 “Wild steelhead SARs from the mid-Columbia River populations … correlated (average r = 0.76) with wild steelhead SARs from the Snake River. Common among these populations (as well as Chinook Salmon PIT tag groups in other regions), SARs were high in 2008 and low in 2011 and 2015.” But it is difficult to go through the chapter and extract all of the necessary information. We suggest that these inter-stock comparisons be extracted to a separate section with some good graphics (e.g. a scatter plot matrix comparing all the stocks) to make it easier for the reader to review and understand. Additional contextual information can be added, i.e. in the above statement, was there anything special about the system in 2008, 2011, and 2015.

The 90% confidence limits look very small in some cases (stock-year) and quite wide in other cases. What are the causes of this; is it simply sample size?

A summary table showing for each species the different variables, years, and stocks would be helpful at the start of the results section. Maybe a time-line plot also to show how much they overlap in time. This could also go into the main body of the CSS.

p.8. Figure A.1 is very helpful. Thank you.

It would be good to include the steps of the calculations and which calculated quantities are shown in figures.

V. ISAB Appendix: Suggested Topics for Further Review 2011-2019

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

Daly, E.A., T.D. Auth, R.D. Brodeur, and W.T. Peterson. 2013. Winter ichthyoplankton biomass as a predictor of early summer prey fields and survival of juvenile salmon in the northern California Current. Marine Ecology Progress Series 484: 203–217.

Faulkner, J.R., Bellerud, B.L., Widener, D.L. and Zabel, R.W. (2019), Associations among fish length, dam passage history, and survival to adulthood in two at-risk species of Pacific salmon. Transactions of the American Fisheries Society 148: 1069-1087. doi:10.1002/tafs.10200

Hayes, M.C., S.P. Rubin, R.R. Reisenbichler, and L.A. Wetzel. 2015. Migratory behavior of Chinook salmon microjacks reared in artificial and natural environments. Journal of Fish and Wildlife Management 6(1): 176–186. doi: https://doi.org/10.3996/022014-JFWM-013

Larsen, D.A., D.L. Harstad, C.R. Strom, M.V. Johnston, C.M. Knudsen, D.E. Fast, T.N. Pearsons, and B.R. Beckman. 2013. Early life history variation in hatchery- and natural-origin spring Chinook salmon in the Yakima River, Washington. Transactions of the American Fisheries Society 142: 540-555. doi:10.1080/00028487.2012.750626

McDonald, LL, and 14 Co-authors. 2007. Study designs for research, monitoring, and evaluation: lessons learned and suggestions for large-scale monitoring programs. Fisheries 32(12):582-590.

Petrosky, C.E., H.A. Schaller, E.S. Tinus, T. Copeland, and A. Storch. 2020. Achieving productivity to recover and restore Columbia River stream-type Chinook Salmon relies on increasing smolt-to-adult survival. North American Journal of Fisheries Management 40:789–803.

Smith, E.P. 2020. Ending reliance on statistical significance will improve environmental inference and communication. Estuaries and Coasts 43: 1-6.

Wilimovsky, N.J. 1990. Misuses of the term “Julian Day.” Transactions of the American Fisheries Society 119 (1): 162–162. https://doi.org/10.1577/1548-8659-119.1.162.