FPC: Our overall conclusion of this review is that Welch et al. (2020) is technically flawed. Therefore, their contentions that Columbia River rebuilding targets may be unachievable, that broad oceanic factors are a primary driving force in Chinook salmon survival coast-wide, and that freshwater management and mitigation activities are misdirected are not supported by the considerable body of available scientific information. Quantitative analysis of their own selected data sets does not support their conclusions. Our primary conclusions are listed below, followed by detailed discussions of each point.

KRS: The CWT dataset is the product of >50 years of collaborative efforts. Below, we respond to all the primary points in the FPC memo. We have also pasted sections from the memo's detailed discussions below their corresponding primary points to marshal related material into the same place.

1. The authors fail to meet the most important requirements of a synthesis or meta-analysis, which is clearly establishing that the data and results utilized are comparable. Failure to understand, acknowledge, and address assumptions, objectives, uncertainty, and details of the data that were used results in comparisons that are uninterpretable. Their approach results in an "apples to oranges" comparison at best, and a misrepresentation of the data utilized and erroneous conclusions at worst.

   KRS: Our goal was not to complete an exhaustive literature review of correlation-based studies linking freshwater events with adult survival, but rather to look at the coastwide patterns of survival. The paper went through two rounds of anonymous peer review, with a total of five peer-reviewers (one reviewer participated in both rounds of reviews). Neither the reviewers nor the editor stated there was a problem with the manuscript meeting the journal's requirements (See Appendix III).

   By requiring that we address all the details in the survival data collected by many government organizations over nearly half a century, the FPC is setting an impossibly high bar—essentially rendering it impossible to synthesize salmon survivals across time or regions. There are too many populations and too many years of data to trace all individual differences. We used the CTC indicator stocks for the main analysis because their CWT-based estimates use a relatively consistent methodology and were collated for international decision-making; they are the best available. The CTC estimates were not provided to us with estimates of error likely because calculation of the real error structure is not currently possible. Additional CWT datasets for the Columbia River Basin were included only after discussion with hatchery biologists. As stated in our paper, we did not calculate SARs directly from the RMIS database because "*we could not verify that adult return numbers from all possible significant components were correctly incorporated and expanded for sampling effort*." We also presented PIT tag-based estimates, but the CWT-based SARS estimates and PIT tag-based SAR estimates were kept separate because of obvious methodological differences.

2. The authors only report results that support their conclusions. They fail to report results of a significant body of analyses that actually evaluate the effects of oceanic and freshwater factors

and the importance of these factors to Chinook salmon survival. The results of these analyses, which are ignored by the authors in their synthesis, conclude that freshwater factors are an important driver of survival for many Chinook salmon populations.

The authors have carefully selected data and analyses in an attempt to support their conclusions. They fail to include a significant body of analyses that actually include ocean conditions, and fresh water conditions and the importance of these variables in resulting survival of chinook salmon. These analyses which are ignored by the authors, conclude that the fresh water life stage environment affects survival to adult.

Important studies on the effects of environmental factors on freshwater survival, ocean survival, and SARs were ignored (Haeseker et al. 2012, McCann et al. 2018, Michel et al. 2015, Cordoleani et al. 2018, Michel 2019). The paper provides no quantitative analysis to refute the results of Haeseker et al. (2012) or McCann et al. (2018) on the importance of environmental factors for explaining patterns in freshwater survival, ocean survival, and SARs. The authors fail to recognize the documented correlations between freshwater and marine survival (Haeseker et al. 2012) or the associations between freshwater environmental conditions and marine survival and SARs (Petrosky and Schaller 2010, Haeseker et al. 2012, Schaller et al. 2014, Michel 2019)

KRS: Our datasets were from reputable sources (documented in SI Table 1) including published papers, the CTC, and the FPC (CSS) itself. Our analytical methods were deliberately minimal, relying on the statistical comparison of median survival levels, in order to minimize the risk of distorting the data. We used all of the data from all of the indicated sources, with the few exceptions documented in the supplementary information section of our paper.

The studies the FPC cites are correlation exercises examining the statistical association of SARs (adult returns) with freshwater conditions when the smolts went to sea. This was not the focus of our study and thus the reason we did not do an exhaustive literature review of this issue, although we did cite two of the listed papers (Michel et al 2019 and McCann et al 2018).

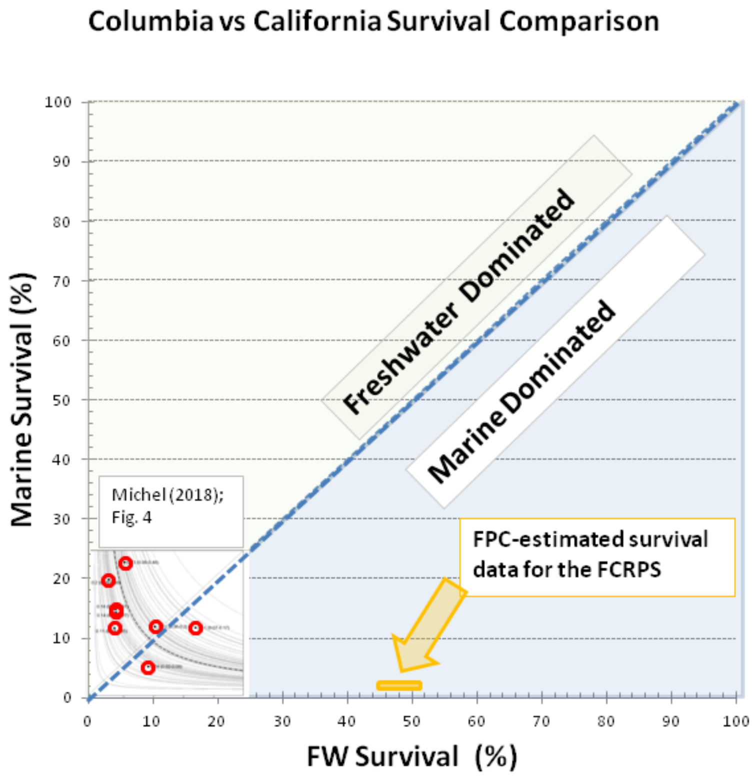
Our view is that the cited studies are incapable of resolving the relative degree that freshwater and marine events control the SAR, because both environments are correlated. At the most basic level, for example, in a La Niña year snowpack is greater and sea temperatures are colder, while in an El Niño year snowpack is lower and sea temperatures are higher. These large-scale shifts change both the timing and amount of freshwater flows and many aspects of the ocean ecology that are only dimly understood.

It is thus not possible to statistically disentangle whether higher flow during outmigration or colder sea temperatures are responsible for higher adult returns without additional information. Correlation studies such as the papers the FPC cite measuring the returning adult numbers cannot do so. This is not an attack on these papers but simply a statement that such exercises cannot resolve whether it is some aspect of flow that drives the relationship or something in the ocean that is doing so.

The Michel et al. (2019, published online in 2018) study the FPC cites is indeed important to the Columbia River issues, but not in the way that the FPC thinks. Michel et al. (2019) reports extremely low freshwater survival for California Chinook; we stated this in our paper: "*The SARs of California Chinook are particularly noteworthy because freshwater survival is exceedingly low*

(Michel et al, 2018); for overall SARS to be higher than Snake River stocks suggests much higher survival during the marine phase".

To clarify this further, we have attached a graph from some of our as-yet unpublished work where we examine the roles of freshwater and marine survival for determining salmon survival. When we compare data for the Columbia and California populations there is far more room for improvement of California SARs. The attached graph compares the two systems with the inset showing the Michel et al. (2019) paper reporting California SARs that both we & the FPC cite. Michel et al. report freshwater survival in the 2~20% range, whereas the FPC's CSS study reports FCRPS survival in the ~50% range. (In the graph the CSS values for FCRPS smolt survival have been multiplied by the below Bonneville smolt survival to the river mouth of ~80% measured by both the JSATs and POST studies to give a more representative comparison).



In the California case, there is far greater scope for freshwater modification to improve adult survival precisely because freshwater survival is already so low (slightly lower than the marine survival; the red dots are scattered near the 1:1 line of equal magnitude). In contrast, the Columbia has far higher freshwater survival than the Sacramento case and much lower marine survival (1/5th to 1/10th the California case). As a result, the scope for significant improvement in Columbia Basin SARs rests on improving survival in the ocean, as doubling freshwater survival to 100% (perfect survival, which would require exterminating all fish-eating bird & resident fish populations as well as removing all 8 dams) can only double the SAR. Thus, unlike the situation reported for California, there is little scope for substantial improvement in adult returns from further increasing freshwater survival in the Columbia unless "delayed mortality" really does exist and is large.

3. There is no doubt that the ocean is important and affects numbers of returning adults. However, the number of smolts that enter the ocean is dependent on freshwater survival and management strategies that result in the highest freshwater survival possible, because not even the best ocean conditions can resurrect a dead fish.

KRS: There is no doubt that fresh water is important, but currently it does not seem to have a large effect on the of proportion of Chinook returning as adults for the majority of populations.

Here is an example:

> Since SARs in the Columbia R area are about 1%, then even a substantial improvement in smolt survival through the FCRPS of, say, Δ=10%, has only a tiny effect on the adult return rate, increasing the SAR from 1.0% to 1.1%. Precisely because hydrosystem survival is already so high ($S_{FCRPS} \geq 50\%$), future SARs can increase to, at best, 2%. Yet achieving this small benefit would require eliminating all sources of mortality in the FCRPS: eliminating all predatory animal life (e.g., all birds and fish), diseases, and direct dam impacts so that FCRPS survival increases from ≥50% to 100%.

Despite the importance of the ocean, we do agree that salmon need good freshwater habitats. In the past, many rivers were degraded by human activities. It took great effort to reverse the damage and continued vigilance is still necessary. Our intention is not to detract from these achievements or to discourage freshwater restoration when it will benefit salmon. However, we are advocating that we approach decision-making in a smarter and more objective way. In the case of the Columbia, those aspects of freshwater habitat affecting survival during the migratory phase (such as the dams) are unlikely to have much influence on determining the poor survival to adult return precisely because other regions have similar survival.

4. Wild spring Chinook stocks in the Columbia River Basin that experience less hydrosystem impacts (e.g., John Day River and Yakima River) have much higher survival than stocks that experience greater hydrosystem impacts (Snake River, upper Columbia River). These examples, which enter the ocean at the same location, clearly demonstrate the impacts of the freshwater environment on Chinook salmon survival.

> *"When the facts change, I change my mind. What do you do?"* John Maynard Keynes

KRS: The FPC are ignoring one of the main points of our paper. In several places, we state that John Day and Yakima "have much higher survival" than Snake River stocks. However, we also pointed out that the FPC's CSS report also estimates PIT tag-based SAR for other populations not migrating through the Snake River dams that have similar survival to the Snake River populations. These additional FPC estimates contradict the FPC's claim that there is delayed mortality.

We also showed that *none* of the CWT-based SAR datasets for the Columbia Basin that we collected to date showed a survival disparity consistent with the delayed mortality theory. Furthermore, two out-of-basin populations (Chilliwack in the Fraser River and an experimental UW hatchery population) also had much higher SARs than other populations in regions where there are no dams. The FPC do not even mention these points, let alone provide a reason why they should be ignored, but are simply reiterating their long-held belief that the two populations they cite fit their preferred theory.

Our observations were thus that (a) dam passage was not necessary for a persistent large disparity in survival to occur between populations, and (b) that of the population pairs that could be compared that did involve Snake River dam passage for one of the pair, all but the two populations the FPC cite *do not* show evidence of the delayed mortality deficit.

Despite having the opportunity to address this point and provide a rational argument for why these counter-examples are wrong, the FPC remain silent about why the other populations in their CSS study do not show the delayed mortality signature. They also have not provided any discussion of why the various CWT-based survival comparisons fail to show signs of delayed mortality for Snake River stocks. (See also our response to some similar claims made in the Schaller et al memo, Appendix II)

5. Historical and recent analyses indicate that freshwater management alternatives that include increases in spill levels and breaching the lower Snake River dams could achieve the NPCC regional SAR goals for Snake River spring-summer Chinook salmon.

"Nothing is so treacherous as the obvious." — Joseph Schumpeter

KRS: Simply put, correlation is not causation. Secondly, the "historical and recent analyses" the FPC refers to often omitted harvest from their correlation exercises assuming it was trivial. However, we found that even for Snake River Spring Chinook (where harvest rates appear to be lowest) the variation in harvest levels over time was large enough to potentially distort statistical analyses using SARs rather than survival. At this point, no one can foretell whether even the sign of these statistical relationships will remain the same once harvests are properly incorporated into studies were based on SARs rather than survival (i.e., where survival is calculated as SARs + harvest).

For example, the important study by Haeseker et al. (2012) examines the effect of spill, water travel time, and hydrosystem survival of smolts on SARs. They concluded that freshwater management alternatives could indeed achieve the NPCC regional SAR goals but they did this by extrapolating well beyond the extent of their independent variables. "Survival" in the Haeseker et al paper is defined for three different life history periods as: "*The life stages that were assessed included freshwater survival (SH, defined as survival from the tailrace of LGR to the tailrace of BON), ocean-adult survival (SOA, defined as survival from the tailrace of BON as a smolt to adult detection at LGR), and life cycle survival (smolt-to-adult survival or SAR, defined as survival from the tailrace of LGR as a smolt to detection as an adult at LGR)*" (p. 123). Neither of the two adult "survival" measures includes harvest before the returning adults reach LGR and are censused, with Haeseker defining the SAR as "*the number of adults that return to LGR divided by the number of smolts that were released at LGR*" (p. 124).

Thus, in Haeseker et al (2012) "*SAR*" really does mean what escapes the fisheries to ***return*** to the dam and harvest is excluded from consideration as a factor influencing "adult return" numbers. This is an important oversight, even for the relatively low levels of harvest we reported in our paper for Snake River spring Chinook.

6. Quantitative analyses of the data sets collated by the authors in this study do not support their conclusions of a coast-wide decline of Chinook salmon survival, that broad oceanic factors are a primary driving force in coast-wide survival of Chinook salmon, or that the SARs of Snake River spring Chinook populations are higher than estimates reported from many other regions of

the west coast lacking dams.

- If Chinook salmon survival across the West Coast is determined in the ocean by common processes, we would expect that common year effects would explain a substantial portion of the variability in SARs. We found that common year effects explained only 14% of the variation in SARs, indicating little support for the claim that common ocean processes are a primary driver of Chinook salmon survival along the West Coast.

KRS: The FPC is taking a simplistic view of oceanic processes. Their analysis assumes that a common process operates on all populations in the same year across the available record. Instead, there appear to be time lags as in the south to north decline in survival that has only recently reached Alaska. Also, a very few populations so far appear to have escaped conservation concerns. Our message is that because survivals are currently similar in most areas including pristine environments, it is difficult to ascribe the cause of the decline to those unique factors whose effects are constrained to specific regions (freshwater habitat degradation in the southern part of the range, and dams in the Columbia River).

- Of the 77 Chinook salmon stocks collated by Welch et al. (2020), 66.2% showed no significant temporal trend in SARs, 14.3% showed significant increases in SARs over time, and 19.5% showed significant decreases in SARs over time. Four of the stocks that showed significant declines over time were hatchery Snake River spring Chinook, wild Snake River spring Chinook, wild Snake River summer Chinook, and wild upper Columbia River summer Chinook.

KRS: There are several reasons why the FPC's trends are not consistent with Kintama's conclusions. First, the FPC takes our conclusions out of context and presents the results on the incorrect scale. We provided a list of the specific <u>regions</u> to which we were referring: "*This [threefold decrease in SARS for hatchery populations] applies to hatchery subyearling Chinook from west coast Vancouver Island, the Strait of Georgia, Puget Sound, and the mid-Columbia River; and to hatchery yearling Chinook from SE Alaska, the lower and upper Columbia River, and the Snake River (upper Columbia and Snake rivers are relative to the historical freeze brand data from Raymond (1988))*." Other regions either have no trend or are generally increasing with time.

Second, the FPC analysis does not match the trends with time visible in the manuscript (our Figure 2) because they fit an exponential curve e^{-bt} to the survival data (or, equivalently, a straight line with slope -b to the log-transformed survival values). This approach is flawed because there is no simple linear pattern between log(SAR) and time (see Figure 2 of our report), violating the basic premise of linear regression that the model is a reasonable fit to the data.

It should not be surprising that such a simple model fit to a wide range of populations failed to show a consistent rate of decrease with time. The log-linear model the FPC chose fails to account for the stanzas of changed marine survival associated with the various recognized ocean regime periods (post-1977: decreased marine survival; post-1989-further decreases and expanded geographic expression; post-1998: improved marine survival; circa 2004: drop to extremely poor marine survival again; post-2008: improved marine survival).

Third, many of the recently developed (and thus shorter) SAR monitoring time series will show an increase in survival with time. This is a consequence of the rapid increase in the number of populations being monitored by various agencies with time, as we showed in Figure 8 of the paper, and some improvement in marine survival in the post-2008 period. The FPC's fitted trends are basically the result of adopting a flawed methodology and uncritically accepting the results because it fits with their wishes, not because it reasonably captures the actual dynamics.

- Welch et al. (2020) claim "most regions of west coast North America with CWT time series extending back prior to the 1978 regime shift show an approximate threefold decrease in SARs for hatchery populations" and that "overall, Chinook salmon survival (SAR) has decreased by roughly the same amount everywhere along the west coast of North America," but did not conduct a quantitative analysis to test these claims.

KRS: Our statement that "*overall, Chinook salmon survival (SAR) has decreased by roughly the same amount everywhere along the west coast of North America,…"* was in reference to the current similarity of survivals in most regions to those in the Snake River (i.e., we were referring to Figure 4 which uses data from 2010-2014 rather than the overall time series in Figure 2). Here we should have written "…has decreased *to* roughly the same amount" as we did at the start of the Discussion and Conclusions sections. We apologize for that error. Our most important finding is that SAR levels are now similar, not that they have declined by the same amount. The statistical analysis is described in section 2.8 in our paper (Comparisons between regions).

- The majority of spring Chinook stocks collated by Welch et al. (2020) demonstrate significantly higher SARs than the Snake River stocks, particularly those hatchery and wild stocks from the southeast Alaska region that are not impacted by dams.

KRS: This is pure fantasy. The FPC claim is based on ignoring our methods. Our analysis used only the most recent years of data (2010-2014) to illustrate the current similarity in coastwide SAR, while their analysis obscures the point by pooling all years together.

The FPC's statement that SARs in Alaska are higher than in the Snake River, is also false. They apparently have made this claim based a comparison of the SARs for wild-origin stocks from southeast Alaska with those of hatchery-origin stocks from the Snake River without telling the reader that this is what they did. (See the column and row titles in the FPC memo Table 1 for the evidence that they pooled hatchery and wild stocks). Thus, the FPC are engaging in exactly the sort of "apples to oranges" comparison that they accuse us of doing. We kept hatchery and wild-origin SARs separate in our analyses.

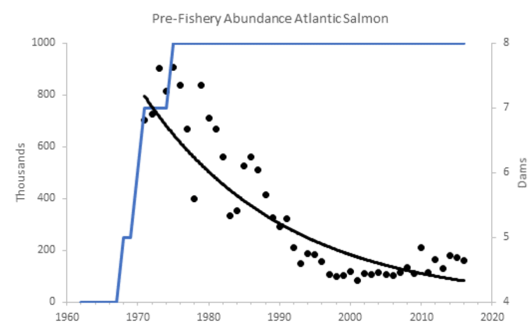
Further, our paper specifically describes and explains the higher SARs of hatchery-origin Alaska stocks relative to Snake River stocks: "*California, northern BC, and SE Alaska yearling SARs were significantly higher than Snake River yearling populations*." We then went on to note that the Alaskan hatcheries used in our analysis "*…are located at sea level and release smolts directly into the ocean after several weeks of seawater acclimation in holding pens, eliminating losses in freshwater*". For wild-origin stocks, we state that "*The median SAR of four wild Alaskan stocks is slightly lower than the median SAR of three Snake River wild stocks when all years of data are considered (Figure 3) and markedly lower when the comparison is restricted to the 2010–2014 time period…*".

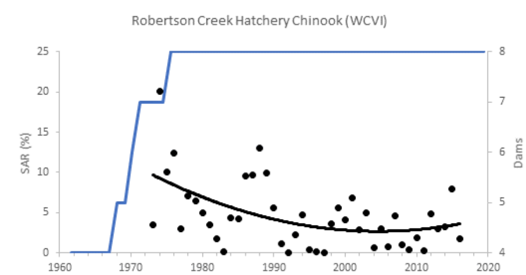
Note that the FPC did not extend their test of our results to subyearling stocks, where SARs for all regions except Oregon were higher than the Snake River region (Figure 4 top panel). A casual reader might mistakenly interpret their point as applying to our results rather than just those selected by the FPC.

- [FPC memo] Figure 3. SARs for Upper Columbia summer Chinook, wild Snake River spring Chinook, hatchery Snake River spring Chinook, and Snake River summer Chinook (points) and fitted regression trend lines that indicated a significant decline in SARs over time. Blue lines represent the number of dams in place during each smolt outmigration year.

"Whenever a theory appears to you as the only possible one, take this as a sign that you have neither understood the theory nor the problem which it was intended to solve."—Karl Popper

KRS: In their memo, the FPC illustrated correlations between dam completion and declining salmon survival as evidence for causation. In the spirit of using this same standard for scientific proof, we have plotted two datasets from outside the Columbia River basin which extend back to the 1970s that show the decrease in salmon SARs for total North Atlantic Salmon abundance and British Columbia WCVI Chinook (Robertson Creek) and the increase in the number of Snake River dams over time. Although these data sets do not extend back to the completion of the earliest Snake River dams, both panels show a decline in SAR/abundance during construction of the two final dams and following completion of the dams. One might conclude, using the FPC's logic, that the construction of the Snake River dams even caused the collapse in salmon productivity beyond the Columbia basin and even in a different ocean.





We have two choices here: we accept the FPC's standard for scientific analysis (correlation), in which case we conclude that reducing the number of dams in the Snake River will also improve the status of a wide range of salmon populations in both the Atlantic and Pacific Oceans, such as those we have plotted. Alternatively, we accept that there are widespread declines in salmon survival, most of which have nothing to do with the number of dams constructed in the Snake River.

Occam's Razor dictates that we should take the most parsimonious scientific explanation, which is that a common driver is primarily responsible for the similar declines in many populations, as we found in our paper. The scientific issue then becomes: "By how much do the Columbia/Snake River dams actually affect Snake River survival separate from climate issues?" These are greatly different perspectives. And to remind the reader of the central premise from our Fish and Fisheries paper, if nearly everywhere is now reporting the same low levels of survival, then a serious re-thinking of current salmon conservation strategies is called for.

7. The authors' comparison of PIT-tag data and CWT data is misleading and reflects the authors' lack of understanding of both types of data and their application.

KRS: We present region-wide CWT-based SARs (Figure 3) and we separately present PIT tag-based SARs for the Columbia River region (Figure 5). The conclusions of our paper are largely based on the CWT data. While we dedicated a section of our paper to comparing CWT-based SAR estimates to PIT tag-based SAR estimates to illustrate the nature of the differences (Figure 6), we do not otherwise draw conclusions based on a comparison of the datasets.

We specifically compared the few available SAR estimates we could find where both stock and outmigration year match across the data types (Fig. 6 of the paper). The slopes of those linear regressions were population-specific but had R^2 values ranging from 87%~99% for subyearling populations and 82-97% for yearling populations. We view these surprisingly high R^2 values as good evidence that both tagging methods give similar survival estimates once harvest and above-dam losses to smolts and adults are taken into account in the PIT-based SAR estimates.

- The author's neglect to address and present the considerable uncertainty associated with CWT SAR estimates. In addition, the authors do not recognize the purpose or the management application of the PIT-tag SARs in the Columbia Basin. Because PIT tags are detected and recorded without the necessity of lethal sampling of the fish, which is required for CWT recoveries, PIT tags are detected multiple times at multiple life stages and provide more extensive life history data and more precise, estimates of SARs.

KRS: We reject the FPC's claim and point first to the three pages (p. 12-14) in our paper where we specifically discuss these issues: Section 4.2 ("*Credibility of SAR Estimates*") and the Conclusions. We used the CTC indicator stocks for our main analysis because their estimates use a relatively consistent methodology and were collated for international decision-making. While imperfect, right now they are the best that are available. CWT-based survival estimates are also the only methodology available for measuring survival outside of the Columbia Basin.

The CTC estimates were not provided to us with estimates of error on the individual SAR estimates, likely because calculation of the real error is not realistically possible. As a result, we used bootstrap resampling of the point estimates to calculate the uncertainty in median SAR levels because if the individual SAR estimates have substantial inherent uncertainty then presumably that additional variability will be captured in the bootstrap resampling of the point estimates. Bootstrapping is an accepted mainstream methodology; for example, Haeseker et al. (2012) employed it and the FPC certainly seem to have embraced its use.

The FPC claims that we do not recognize the purpose or the management application of the PIT-tag SARs in the Columbia Basin. In fact, one of our important conclusions was that the exclusion of harvest may be compromising the use of PIT tag-based SAR estimates underpinning NWPCC recovery goals. We have no reason to question the relevance of PIT tags for measuring *smolt* survival in the hydrosystem, but their use to measure adult survival is fundamentally compromised because the PIT tag system excludes harvest. Nearly a quarter century of interpreting SARs as "*survival*" instead of "*what is left over from the fisheries*" went unrecognized.

In summary, the FPC are essentially saying that it is not possible to use the CWT-based SAR estimates because there is too much uncertainty in the data. Our position is that the CWT dataset has weaknesses (which we discuss in the paper) but that it is also an important scientific resource. We certainly need to adapt as and when we get better information. Neither PIT tags nor CWT tags are perfect. We used both tag types to demonstrate that (a) survival has fallen to similar levels, (b) concerns with using PIT tags as measures of adult survival—but probably not smolt survival—are serious, and (c) CWT and PIT tag-based survival estimates track each other surprisingly well at the level of individual populations. The deviations from a perfect 1:1 relationship seem to be primarily in aspects of survival that PIT tag-based estimates currently fail to track.

8. The authors have misrepresented Pacific Salmon Commission and Chinook Technical Committee (CTC) data sources. The CTC does not maintain a smolt-to-adult-return database, contrary to the authors' statements.

    - CTC does not maintain a database of smolt-to-adult return rates based on CWT data. The attribution of the CWT SARs to the CTC is not accurate. Furthermore, the CTC chairperson advised that the authors assigned incorrect smolt-ages for three stocks (NSF, SKF and SQP). The CTC expressed serious concerns with how the CTC data were characterized in the Welch et al. (2020) paper. The CTC uses and maintains various data sources that contain the information necessary to calculate the SARs calculated in Welch et al 2020, as they are defined in the paper, but they were not developed for this purpose. According to the CTC chair, no current CTC members were afforded the opportunity to review Welch et al.'s 2020 manuscript prior to publication.

    - Welch et al. (2020) define the smolt to adult return rate (SAR) as "the threefold product of freshwater smolt survival during downstream migration multiplied by the marine survival experienced over two to three years in the ocean and multiplied by adult freshwater survival during the upstream migration to the final census point." It is important to recognize that this definition of the SAR includes the combined impacts of freshwater and marine survival factors, and does not separate the effects of the two environments. Therefore, the characterization that SARs only reflect marine processes is a mischaracterization. The CWT recovery data used by the Pacific Salmon Commission (PSC) Chinook Technical Committee (CTC) are not compatible with this definition of SARs and the CTC does not use their CWT recovery data to calculate SARs. CWT recoveries in ocean fisheries overestimate survival to adult return because these data represent harvest mortality (not survival to adult return in terminal areas) and the natural mortality that occurs in the marine environment prior to adult return in freshwater is not incorporated. CWT recoveries in freshwater also likely underestimate survival to adult return because straying and enumeration of hatchery fish on spawning grounds is often incomplete.

KRS: The FPC suggests that we were misleading about the source of our CWT SARs estimates. After reading the FPC's claims about the CTC concerns, we emailed the two current co-chairs of the CTC (John Carlile & Antonio Velez-Espino) asking for clarification. The email trail with the current CTC chairs (with our summary and analysis of what measure of adult abundance is appropriate for our analysis) is included as Appendix IV.

The SAR data we attribute to the CTC were sent to us by Dr. Gayle Brown, the Canadian Co-Chair of the CTC until her retirement in July of 2020[1]. Gayle retired after we had submitted our paper. She also met with us multiple times to explain its use and clarify our questions. Dr. Kristen Ryding, who is a US member of the CTC, also assisted us and provided us with the formula that had been used to calculate the estimates. Neither were asked to review the manuscript as we expected that would be done by the journal's reviewers.

Despite these conversations, we were unaware that the CTC did not consider the SARs estimates we presented as "official products". However, they were calculated using a transparent formula (section 2.2 in the paper) from vetted CTC data sources. As per the CTC response to our request for clarification on what the FPC is claiming (Appendix IV), the CTC "*uses and maintain various data sources that contain the information necessary to calculate SARs as they are defined in the paper*". The FPC are making a mountain out of a molehill.

The email trail with the current CTC chairs (with our summary and analysis of what measure of adult abundance is appropriate for our analysis) is included as Appendix III of this document and on our website. As for the other point about "not consulting" with current members of the CTC, the SAR data we attribute to the CTC were sent to us by Dr. Gayle Brown, the Canadian Co-Chair of the CTC until her retirement in July of 2020[2]. Gayle retired after we had submitted our paper. She also met with us multiple times to explain its use and clarify our questions. Dr. Kristen Ryding, who is a US member of the CTC, also assisted us and provided us with the formula that had been used to calculate the estimates. Neither were asked to review the manuscript as we expected that would be done by the journal's reviewers.

We used the measure of SAR we chose because it is calculated similar to our other CWT data sources (sections 2.3 Agency estimates and 2.4 PSMFC estimates), with the exception that it is adjusted for incidental mortality from harvest. It is also similar to our definition of SAR as the "*threefold product of freshwater smolt survival, marine survival, and adult freshwater survival"*. The alternative measure from our CTC data source was the CTC's brood year survival rate estimates from release until entry to the fishery at age 2 or 3 (section 2.1.3 of TCCHINOOK (19)-02) that are calculated for the purposes of harvest management. These values are a "formal product" of the CTC and estimate the ocean abundance of young salmon prior to the start of the fisheries. See our direct response to the CTC (Appendix IV), where we compare the measure of adult returns we used with the abundance of immature Chinook in the ocean prior to the start of fisheries, the "pre-fishery" abundance the CTC consider an official product; there is a strong relationship between the two estimates of abundance. We believe that the measure we used is more appropriate than an estimate of abundance prior to fishing.

---

[1] The Acknowledgements section of our paper states (in part): "We particularly thank **Dr Gayle Brown** (DFO; retired) for providing access to the Chinook Technical Committee's SAR database and for many discussions clarifying the interpretation and use of the data. We also received significant assistance in understanding critical details of many SAR and harvest data sets from scientists from *[deleted]* ...**WDFW (Kristen Ryding)…**". Both Brown & Ryding were members of the CTC. The FPC were presumably aware of this when they worded their criticism.

We never stated that *"SARs only reflect marine processes"*. In fact, we explicitly said the opposite (see the quote above about the SAR being a "three-fold product"). What we did say—and still stand by—is that if regions with excellent freshwater habitat have the same low SARS as regions with compromised freshwater habitat then the common problem must be in the ocean, not freshwater. See the expanded analysis we report below of Alaska hatchery SARs relative to the Snake River region for fresh evidence that the decline in SARs really is occurring in the ocean and not in freshwater.

Finally, the FPC claims that we assigned incorrect smolt-ages for Nooksack spring fingerling, Skagit spring fingerling, and Squaxin pens fall yearling Chinook. The age at migration as yearling/subyearling was correctly assigned according to the CTC reports. Subsequent discussion with the CTC clarified the source of the error, which is that the year of outmigration is offset by +1 for the two spring stocks and -1 for the fall stock. The data were sent to us with this error. Fortunately, it makes no material difference to our results and does not affect our conclusions (Appendix IV).

9. The authors conflate fall Chinook and spring Chinook and misrepresent the Northwest Power and Conservation Council (NPCC) regional 4% average, 2% to 6% SAR goals. The authors ignore the fact that these two different races of Chinook salmon have different life histories and have a very different presence in ocean fisheries.

    - While Welch et al (2020) referenced early work by Marmorek et al. (1998) regarding the 2-6% SAR goal they failed to take into account multiple years of analyses on Snake River populations published in CSS reports such as Chapter 5 of McCann et al. (2018). In those reports, the CSS demonstrates the applicability of those SAR targets to the Snake River wild spring/summer Chinook and steelhead. The SAR goal applies to populations in a very specific life segment and takes into account population productivity in spawning areas. A SAR target appropriate for Snake River spring summer Chinook salmon, measured through a very specific life phase, was not developed for, and is not appropriately applied to, CWT populations up and down the Pacific Coast. It is extremely difficult to match the exact life stage and mortality that populations would have occurred prior to entry into that life stage and subsequent to it.

    - Mortality prior to the upper dam and terminal fisheries are different throughout the various river systems.

    - Welch et al. (2020) make assumptions about SAR goals and then make conclusions based on a lack of information. For example, they state in their discussion section regarding the importance of ocean harvest "Unfortunately, what went unrecognized was the effect on the many Columbia River studies based on PIT tags." However, the CSS, recognizing the importance of ocean harvest for fall Chinook salmon, does not consider the 2-6% SAR goal appropriate for fall Chinook. This lack of understanding of CSS PIT-tag analyses makes these broad concluding statements nonsensical.

KRS: The FPC has worded this point to suggest that we ignored the life history differences between fall and spring Chinook. Instead, what they mean is that we were not sufficiently clear on how we used the 2-6% recovery goals.

We split our analyses into separate sections for fall and spring Chinook. We can find nowhere in our published paper where we "conflate" the results as the FPC claims. Section 2.7 ("Division by Life History") explains at length why we kept the two major life history types separate, but to quote briefly from the paper: "*These life history types are examined separately in our analysis because there are important ecological differences between them*" (p.198).

As for the NPCC 2-6% SAR goals, we have shown these as bands on all regions and both major life history types, not because we wish to misrepresent the applicability of the NPCC's goals but because they represent the only explicit SAR recovery targets for the west coast of North America. We showed these recovery goals as bands on all the regional fall Chinook SAR estimates as well as the spring Chinook regions (Figs. 2, 3, & 5) because they are relevant to the debate about whether they can realistically be attained. The text makes repeated mention of how the recovery targets apply to spring Chinook. Because the various plots show that these goals are not being reached in other regions, then surely the debate in the Columbia should begin to take into account the point that other regions without dams are not achieving these goals either... including the more productive fall Chinook stocks.

The FPC say it is "nonsensical" of us to conclude that exclusion of harvest may affect Columbia River studies based on PIT tags. We disagree. Our conclusion about the effect of harvest was not so much related to the 2-6% objectives but based on our concerns that the unexpectedly high and variable levels of harvest over time plus the Pacific Salmon Treaty's abundance-based harvest management system may set up a negative feedback system that distorts the adult return levels in complex and important ways. In combination, these two concerns could partly or completely mask fluctuations in the number of returning adults, which is the current metric that the Columbia River Basin sets for recovery. (Incidentally, in-river harvest management in the Columbia River could potentially also lead to the same sort of feedback system, not just the Pacific Salmon Treaty).

Unless the issues concerning the failure to incorporate adult returns can be sorted out it may be more realistic to set smolt survival goals under the ESA relative to hydrosystem operations than make hydrosystem operations dependent on what escapes from unmonitored fisheries.

10. The analytical techniques and underlying data used by the authors are inappropriate, misleading, and unreliable.

- The sample sizes in these groups are concerningly unbalanced and small varying from 0 to 11 (Table 2) and from 3 to 49 (Table 3), despite the fact that these groups were visually weighted equally with uniform confidence estimates.

- The replication process relied on unbalanced sample sizes between years which skew the expected draws within regions and weights SARs to years that are more heavily represented within the sampling distribution. For example, drawing at random from the 2010-2014 SAR distribution for the Snake River (N=46), a SAR value from ocean entry year 2012 would be 5.5 times more likely to be drawn than an SAR value from 2014.

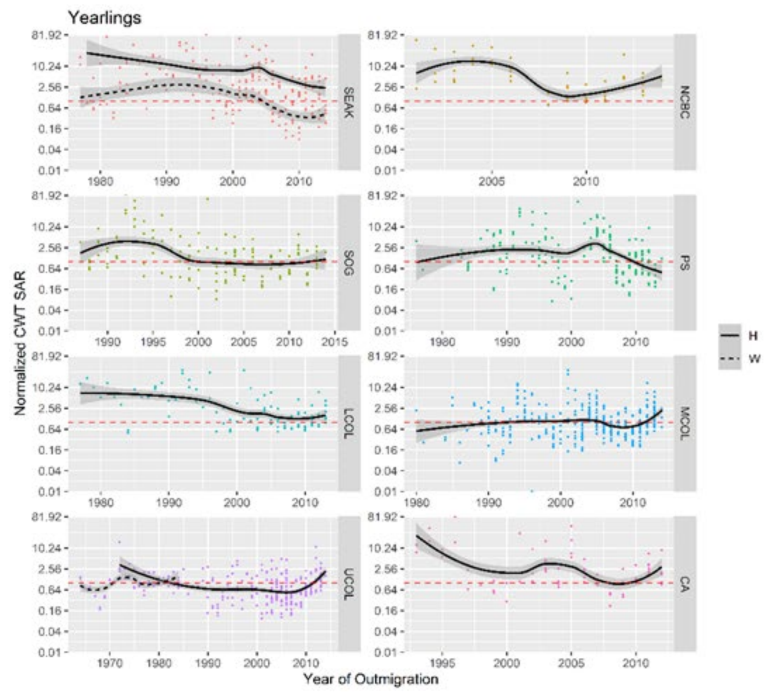
- The lack of any attempt to include or evaluate a year effect within the Welch et al. (2020) analysis significantly dilutes the level of insight we can obtain from the visual comparisons presented in their Figure 4.

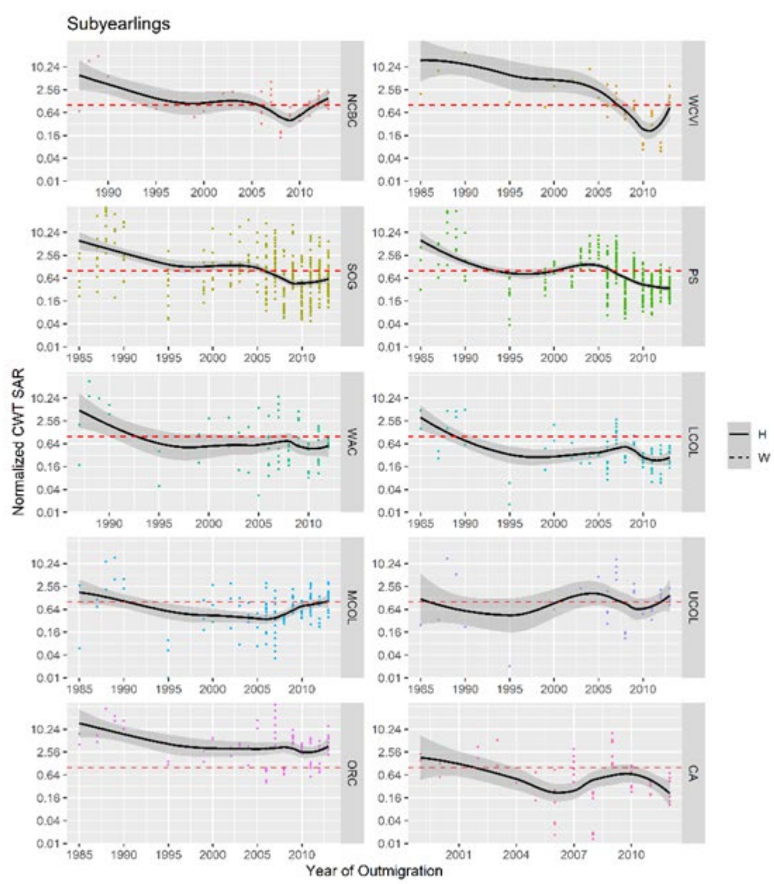
KRS: The five reviewers and the journal editor did not find that our techniques and data were inappropriate.

Regarding sample size, the CWT dataset is what it is: the number of populations monitored in each region varies. It is reasonable to assume that the CTC has enough indicator stocks to adequately monitor each region. More is always better and balanced designs are better, but we work with what is available.

Our original analysis was simply focussed on representing 'recent' survival; within the 5-year period 2010-2014 (Fig. 4 of the paper), we did not view it as particularly important if more samples came from certain years. Here the FPC raise an interesting point: the imbalance between years in the number of SAR estimates could potentially distort the results if some years have better ocean survival and not all regions have sufficient data submitted to properly capture this.

To examine the FPC's criticism, we redid our analysis using all years in the time series but this time restricting the comparison to individual years. The number of datapoints available for each year was small, so instead of bootstrap resampling, as we did in the paper, we calculated the annual SAR ratio $SAR_{i,t}/SAR_{Snake,t}$ for each of the i regions, with each year t treated separately for all possible combinations of $SAR_{i,t}$ and $SAR_{Snake,t}$. (In other word, if two regions have three SAR estimates each in the same year, (a,b,c) and (f,g,h), we calculated





all of the possible unique ratios: (a/f, a/g, a/h, b/f, b/g, b/h, c/f, c/g, c/h)).

The figures to the right use a separate panel for each regional comparison with the Snake River. We used the standard loess smoother in R to highlight the trend over time in the SAR ratios (hatchery populations: solid black line; wild populations: dashed black line). The grey bands show the 90% confidence intervals of the smoothed average; they do not include any underlying error in the SAR estimates themselves because CWT-based SAR estimates do not include these values, so it is the variability among the annual point estimates of the SARs for a region that is used to define the uncertainty.

We specified the 90% confidence interval because the two limits are the appropriate test statistic for a one-sided test at the 95% confidence level of the hypothesis that the i^{th} region's SAR values include the Snake River level. The horizontal dashed red line shows the null hypothesis of equal SARs with Snake River values (i.e., $SAR_i / SAR_{Snake} = 1$). Consistent with our 5 year analysis, all regions except for north-central BC and SE Alaska hatchery yearling populations and Oregon coast subyearlings have median SARs trending down over time either reaching or falling below Snake River values. (Note that the y-axis scale is logarithmic and that the span of years on each x-axis changes).

As described in our paper, the five SE Alaska hatchery populations are reared at sea level and only released after a period of acclimation to seawater in netpens; as a result, SE Alaska hatchery SARs do not include any component of freshwater smolt survival. Although the higher SAR values of Alaskan hatchery Chinook relative to the Snake River needs to be qualified by this caveat, the downwards trend to more similar SAR ratios with time is clear.

This approach extends the SAR comparison from the recent 2010-2014 aggregate period reported in our paper to all years of available data with each year treated individually. Our original conclusion for 2010-2014 thus stands, but the longer time series also shows that other regions were formerly far more productive than they are now (SAR ratios relative to the Snake River $\gg$1) and that relative productivity fell over time to levels at or below the numerical SARs reported for the Snake River in many regions. Even hatchery SARs for SE Alaska, which exclude migration losses in freshwater, show a large reduction over time in survival relative to the Snake River, and this reduction parallels the decline in wild Alaskan SARs over time.

This is additional evidence that the survival decline occurs in the marine phase because the rearing strategy for the Alaskan hatchery populations excludes exposure to freshwater in the smolt phase.

References

Haeseker, S. L., J. A. McCann, J. Tuomikoski and B. Chockley (2012). "Assessing Freshwater and Marine Environmental Influences on Life-Stage-Specific Survival Rates of Snake River Spring–Summer Chinook Salmon and Steelhead." Transactions of the American Fisheries Society **141**(1): 121-138 DOI: 10.1080/00028487.2011.652009.