Review of: Welch DW, Porter AD, Rechisky EL. A synthesis of the coast-wide decline in survival of West Coast Chinook Salmon (Oncorhynchus tshawytscha, Salmonidae). Fish Fish. 2020; 00:1–18. https://doi. org/10.1111/faf.12514

Review by: Howard Schaller, Ph.D. Retired Project Leader of the Columbia River Fish and Wildlife Conservation Office, US Fish and Wildlife Service; Charles Petrosky, Ph.D. Retired Fisheries Biologist, Idaho Department of Fish and Game; and Margaret Filardo, Ph.D. Retired Senior Fisheries Biologist, Fish Passage Center

November 23, 2020

"The lot of critics is to be remembered by what they failed to understand" -George Moore

Kintama note to readers: We have chosen to respond point by point to Schaller et al.'s detailed critique, which followed the "Summary points" below and begin on page 4.

As an overview, we disagree with Schaller et al. for two main reasons. We acknowledge and respect the authors as solid scientists, but they put too much faith in (a) statistical exercises correlating adult returns (not survival) from the fisheries (SARs) and (b) the use of simplifying assumptions about where salmon survival is determined. We now know that (a) has serious flaws (see our paper). The critical period concept (b), has been a simplifying assumption used by fisheries scientists since Hjort first proposed it in 1914, but has never been proven. It is unwise to predicate the success of multi-billion dollar salmon conservation efforts on an unproven hypothesis.

Summary points

(Note: Three pages of summary points by Schaller et al have been deleted because we respond to the detailed "Specific Comments" section below, which expands on each of the summary points).

Specific comments (by section)

Abstract

The authors state "Given the seemingly congruent decline in SARs to similar levels, the notion that contemporary survival is primarily driven by broader oceanic factors rather than local factors should be considered." In fact, numerous studies have directly considered local and oceanic factors for PIT tag SARs and life cycle survival rates (which the authors don't cite). These studies have shown that both freshwater migration conditions and marine conditions are highly influential to survival (Petrosky and Schaller 2010, Haeseker et al. 2012, Schaller et al. 2014, Petrosky et al. 2020).

The authors state: "Ambitious Columbia River rebuilding targets may be unachievable because other regions with nearly pristine freshwater conditions, such as SE Alaska and northern BC, also largely fail to reach these levels." This authors' conclusion was based primarily on hatchery stocks from these regions, which are irrelevant to wild Interior Columbia stream-type Chinook SAR goals.

The authors claim that PIT tag SAR estimates from the Columbia River Basin are generally consistent with CWT findings is a questionable conclusion. Their primary analysis is based on five years of data, with limited environmental contrast. Studies that have found that both freshwater and oceanic factors are influential to salmon survival are based on long time series, encompassing a large degree of variation in environmental and management conditions (Petrosky and Schaller 2010, Haeseker et al. 2012, Schaller et al. 2014, Petrosky et al. 2020).

The authors statement that PIT tag-based SARs are not adjusted for harvest which compromises their intended use, is extremely misleading. The authors fail to point out the several advantages that PIT tags have over CWT, which are especially important in systems with dams and allow for detecting tags in ESA listed fish without having to sacrifice fish to recover a CWT. Including harvest in PIT tag SARs is a relatively simple matter for Interior Columbia stream-type Chinook. In addition, many published analyses have incorporated harvest into PIT tag-based SARs (Schaller et al. 2007, Petrosky and Schaller 2010, McCann et al. 2017, McCann 2018, and Petrosky et al. 2020). These studies illustrate different temporal and spatial patterns in survival rates and document a strong influence of both freshwater and marine conditions on SARs and life cycle survival rates.

KRS: Our paper extends beyond the Columbia River basin. We cited Haeseker et al. (2012). The Petrosky et al. (2020) paper was published nearly concurrently with ours.

Our conclusion regarding rebuilding targets is based on all data presented in Figure 3 of our paper- hatchery, wild, fall, and spring Chinook (treated separately). Surely Schaller et al should be concerned that wild spring Chinook from Alaska have mean/median SARs below that of wild Snake River spring Chinook, as well as all of the hatchery comparisons that showed similar patterns? It is all very well to talk about the complexity of the CWT data, but neither Schaller et al nor the FPC in their critiques have provided any mechanism that would explain why decades ago other regions had higher SARS than the Snake River (which everyone cheerfully accepts) but these regions then drop to equivalent to or lower survivals than recent Snake River levels.

The high R^2 values (between 0.82-0.99) we found between CWT and PIT tag-based SARs for individual populations are an important result. The high correlation should be viewed as indicating that the two survival monitoring technologies track each other well. If Schaller et al would like to present actual reasons to back up their claim that this result is "questionable", we would be happy to consider it. In any case, their stated rationale that we rely on "…five years of data, with limited environmental contrast" (Fig. 4) is incorrect; the CWT-PIT tag comparison was based on all years of data we could find "…*where both tagging methodologies were employed in the same year*" (Fig. 6).

We recognize that harvest is incorporated into several peer-reviewed publications cited by Schaller et al and is included in the CSS's lifecycle model for Snake River spring Chinook; however, the main product of the CSS report, SAR for numerous populations of Chinook, do not include harvest. According to McCann et al (2018): "*The NPCC (2009 and 2014) SAR objectives did not specify the points in the life cycle where Chinook smolt and adult numbers should be estimated. However, the original PATH analysis for Snake River spring/summer Chinook was based on SARs calculated as adult and jack returns to the uppermost dam (Marmorek et al. 1998)….We have made preliminary comparisons of the overall SAR estimates for wild groups to the NPCC 2%–6% SAR objectives, recognizing additional accounting for harvest, straying and other upstream passage losses may be needed in the future as NPCC and other SAR objectives are clarified.*"

Thus, if the NPCC SAR goals are based on CSS SAR estimates features in their reports, harvest should be included. According to NOAA's Recovery Plan for Snake River Spring/Summer Chinook (NOAA 2017): "*Harvest exploitation rates have been relatively low on Snake River spring and summer Chinook salmon, generally below 10 percent, but have increased in recent years due to the continued large returns of hatchery spring Chinook salmon to the Columbia River basin. These large returns triggered increased allowable harvest rates under the abundance-driven sliding-scale harvest rate strategy guiding annual fishery management.*" Our view is that ignoring harvest is a serious omission in studies examining how SARs are impacted by hydropower operations. While not all studies make this error, many do.

1. Introduction

Welch et al. sets up a false dichotomy of ocean versus freshwater influences on salmon survival in literature (whereas a wealth of existing literature shows it can be both in altered systems).

Not sure they accurately reflected NPCC 2-6% SAR goals; they do not acknowledge the parallel objective of understanding ocean influence so that actions taken in freshwater can help ensure salmon can survive in face of varying ocean conditions.

The introduction seems to be setting up a means to discredit NPCC SAR goals for wild interior Columbia stream-type Chinook, but they then focus most of the analysis on hatchery ocean-type Chinook. Recent analysis in Petrosky et al. (2020) gives more direct evidence supporting NPCC goal relative to Marmorek et al (1998) and NMFS Interior Col. River rebuilding goals (ICRTRT 2007). This support is for stream-type Chinook and by analogy steelhead. It is well understood that 2-6%SAR goal applies to stream-type Chinook and steelhead given similar life history expressions. One wouldn't expect ocean type fish that emigrate shortly after hatching to have similar goals. This is complicated by their analysis which relies on hatchery ocean type Chinook, which have nothing to do with wild stream-type Chinook.

Welch et al. misleadingly states that harvest is not included in PIT tag-based survival estimates. They further state that the "*previously unrecognized limitation* of PIT tagging methodologies is critical to current conservation efforts in the Columbia River Basin because of changes to the terms of the US-Canada Pacific Salmon Treaty…" This is completely false; users of the PIT tag data are fully aware of the portions of the life cycle incorporated in PIT tag estimates. Note that it is a trivial matter to incorporate harvest effects for Snake/Columbia stream-type Chinook populations, which are not intercepted in ocean fisheries. In fact, harvest has been accounted for in many published analyses using PIT tags (which they neglect to cite).

The authors describe a 3-fold process for calculating CWT based SARs (which they don't apply), then they criticize PIT tag SAR methods that doesn't include animals removed by fisheries. However, this statement is false, because in many instances PIT tag SAR methods do include fishery pacts (Petrosky and Schaller 2010, McCann et al. 2017, McCann et al. 2018, and Petrosky et al. 2020.

At broadest level of stock comparison, they claim most of the population decline is due to the ocean via common processes. Their broadest level comparison of populations ignores the different life-history types for these stocks. They don't do any direct analysis of how ocean conditions influence patterns for SARs (the authors comparison employs a very short time series of SARs). The authors conclude that ocean conditions drive these coastwide SAR patterns and freshwater conditions have limited influence on SARs. The authors ignore numerous published studies that have directly considered local and oceanic factors for PIT tag SARs and life cycle survival rates. These studies have shown that both freshwater migration conditions and marine conditions are highly influential on Chinook survival rates (Petrosky and Schaller 2010, Haeseker et al. 2012, Schaller et al. 2014, Petrosky et al. 2020). For these reasons, the broad generalization (by the authors) that most salmon conservation problems are caused by ocean conditions is extremely misleading.

KRS: Many of the points raised here are discussed in subsequent sections.

2.1 Methods/Data Sources

The authors analysis heavily relies on hatchery subyearling Chinook groups. The comparison of survival estimates of these subyearling stocks to SARs for Interior Columbia stream-type Chinook is questionable, because of the difference in life history patterns and period of ocean residence. One may expect to see some general relation of patterns of survival due to large scale ocean conditions, but the authors' approach looked at very short time series of SARs, which would preclude looking at temporal SAR patterns in relation to variable ocean conditions (given past decadal scale oscillation of ocean conditions, Mantua et al. 1997) . However, with the very different life histories, tagging locations and ocean entry points, it's a stretch to expect magnitude of SARs to be similar.

KRS: Schaller et al's argument that we relied "heavily" on hatchery subyearlings is untrue. Although there are more subyearling datasets (n=50 total), the number of yearling datasets (n=33 total) was reasonable (Fig. 3). We kept subyearling and yearling analysis separate (see Section 2.7 Division by life history). It would be nice if the time series were longer, but we used all available data.

Authors claim PIT tags transmit a signal. PIT tags don't transmit – they're passive receptors after being energized thru antennae array to transmit. The large infrastructure of detection arrays in the Columbia River located at dams and instream sites is why they're used here. PIT tags are more accurate and precise for survival estimates. Using PIT tags, one can differentiate migratory pathways thru system.

With CWTs you need to kill a fish to recover the coded information, whereas PIT tags allow for multiple mark/recapture methods. Authors never present or use the large uncertainty surrounding CWT based survival estimates. The authors fail to mention any short comings of CWT resulting in a very one-sided presentation. Given the numerous ESA listed fish populations, PIT tags are a more conservation-oriented method since fish don't need to be killed to read a tag code.

KRS:

1) Schaller et al have the technology backwards. PIT tags do transmit the signal. When a PIT tag enters the EMF field generated by a PIT tag receiver's antenna, an internal copper coil is energized, providing sufficient electrical power for the tag to transmit its unique tag code back to the receiving antenna. Apart from tapping into the electromagnetic field, PIT tags have no technology capable of reception once manufactured—they simply blindly transmit their code when they have the power to do so.

2) "*PIT tags are more accurate and precise for survival estimates*". This is an important but false claim. Schaller et al.'s statement makes it clear that the full implications of the technical problems we identified with PIT tags are not yet understood[1]. If there was no harvest, Schaller et al.'s claim that survival estimates using PIT tags are more accurate and precise than CWTs would be correct—but there is harvest. We will formally spell out why Schaller et al. are wrong.

We define accuracy and precision using the standard statistical conventions. An accurate estimate is one with no bias, so that with sufficient sampling the estimate converges on the true value. A precise estimate is one with small variance. In the current context, the estimate of interest is the proportion surviving, S, and the estimated variance on the binomial proportion S is $\sigma^2 = N \cdot S \cdot (1\text{-}S)$, where N is the sample size. The standard error on the proportion is then $SE=\sigma/\sqrt{N}$, so an approximate 95% confidence interval on S would be S±2SE.

In the absence of losses to harvest, survival would indeed equal the PIT tag-based SAR, SAR_{PIT}=Adults Returning/Smolts Released. The PIT-based estimate would then be *accurate* because harvested fish do not have to be accounted for. In practice, what the PIT tag system measures is the number of adults *surviving the fisheries to return* to the Columbia River dams and detected there. This is not survival but rather the escapement from the fisheries; the two are equal only if the harvest is either zero or is properly accounted for and added to the number of adults surviving to return to the dams and counted by the PIT tag receivers. Since PIT tag-based estimates do not account for harvest, Schaller et al.'s first claim that the PIT tag system is more accurate is false—it is a downwards biased estimate of survival and the degree of this bias is potentially large and varying with both time and between populations—at least for those populations for which we found data (see Fig. 8 of our paper).

Harvest needs to be added to what is censused at the dams by the PIT tag detectors. Although the degree of bias depends on the population, in the cases where we found harvest data the distortions caused by this bias *were large when compared with the expected effect of dam manipulations* on the number of subsequently returning adult (total returns, catch plus escapement) that scientists are trying to detect. Thus PIT tag-based estimates are generally not more accurate than the CWT-based estimates but less so because CWT-based estimates include harvest, even if the estimate is imperfect.

We will now explicitly demonstrate that Schaller et al.'s claim that PIT tag-based SAR estimates are more *precise* than CWTs is also probably wrong because the PIT tag-based SAR estimates must suffer from the same issues that they list for CWT-based SAR estimates.

The standard error on a binomial proportion, *p*, is simply $SE(p)=\sqrt{p(1-p)/N}$, where N is the sample size (# of smolts released in the case of the SAR). This is the "ideal" precision that can only be obtained when both the numerator and denominator of the survival proportion are known without error.

[1] We leave aside for the time being the equally (and possibly more) serious issue of the Pacific Salmon Treaty acting as a negative feedback loop distorting the adult return. We address that issue elsewhere.

In our paper, we reported that the fraction of the adult return that was harvested, h, was large and variable. We also showed that to turn the PIT tag-based SAR estimate into an estimate of survival required dividing through by the fraction not harvested, (1-h), so that

$$S_t = \frac{SAR_t}{(1-h_t)}.$$

(we disregard the complexities of the multiple age-structure and multiple ages at return and harvest to focus on the key issues). Survival corrected for the missing harvest is thus the ratio A/B, where A is the proportion of PIT tagged adults surviving to be censused at the dams (i.e., the currently used PIT Tag estimate, SAR_t), and B is the proportion of unharvested PIT tagged adults, ($1-h_t$). A standard result from statistics is that the first-order (Taylor series) approximation to the variance of a ratio A/B is:

$$Var(\frac{A}{B}) = \frac{A^2}{B^2}\left\{\frac{Var(A)}{A^2} + \frac{Var(B)}{B^2} - \frac{2\Box Cov(A,B)}{A\Box B}\right\}$$

Covariance of the SAR estimates and the fraction of adults unharvested is probably small, so we can reasonably ignore the third term. We can then re-write this equation in more concrete terms as follows[2]:

$$Var(\frac{SAR}{1-h}) \approx \frac{SAR^2}{(1-h)^2}\left\{\frac{Var(SAR)}{SAR^2} + \frac{Var(1-h)}{(1-h)^2}\right\}.$$

Only if harvest is zero does the variance on the estimate of survival collapse to the PIT tag-based variance term, Var(SAR)= $\sigma = (N\cdot S\cdot (1-S))^{-2}$. When harvest is positive, the true variance is larger. The first term outside the braces is an inflation factor $(SAR/(1-h))^2$. If, for example, harvest is 50%, this factor is 4X larger than the variance would otherwise be. Even in the best-case situation of the Snake River Spring Chinook where reported harvests are lowest, we found harvests varied from h=10%~25%, this inflation factor varies from $1/(1-h)^2$=1.25X~1.75X. This term alone means that the uncertainty in PIT Tag derived survival estimate is substantially larger than generally believed; rather than the standard large-sample assumption that a 95% confidence interval on a SAR estimate should be ±2SE, this term increases them to ±2.5 SE to ±3.5 SE for Snake River Spring Chinook in the last two decades; for Fall Chinook the increase is from ±2 SE to ±8 SE.

However, to this increase we still need to include the second term within the curly braces, $(Var(1-h)/(1-h)^2)$. As a result, the variance on an estimate of survival has to be larger than the variance on the SAR. How much larger? Although it is generally not explicitly discussed, almost all estimates of both harvest and harvest rate depend on allocating the catch from the various fisheries to specific populations… and that is generally done based on the CWT-based sampling of the harvest in the various fisheries. So, whatever the unknown variance structure of CWT-based SARs is,

[2] We thank Dr John Skalski for a discussion on the finer points of this derivation and for pointing out an error in our original starting point for the development.

Schaller et al.'s complaint about them being poorly defined also applies to the harvest.

In other words, despite Schaller et al.'s claims that PIT tag-based SAR estimation is more accurate and precise, it is neither—and it cannot be so long as fisheries intercept even moderate amounts of the salmon return and the degree of harvest is uncertain. So, whatever the concerns are about CWT-based survival estimates (and we acknowledge there are important concerns), those concerns also apply to a degree to the PIT tag-based survival estimates—and the degree that it applies depends on the (a) the harvest level and (b) the degree that harvest is determined using CWTs.

Given the way the Columbia River PIT tag system was set up to study smolt-to-adult survival, it cannot be otherwise. To avoid the problems with using CWT estimates will require using a separate source of more reliable information on how many Columbia River salmon of a particular population group were harvested in all the west coast fisheries than CWTs provide—the PIT tag system abdicated responsibility for providing this information to the CWT system decades ago when the choice was made not to survey the sport and commercial harvest for PIT tags.

In reality, the return of adults that escape from the various fisheries (the SAR as measured using PIT tags) *is not the biological question of interest* from a policy standpoint—it is the fraction of smolts surviving to adulthood. The allocation issue of distributing adult salmon to various fisheries versus leaving them to escape back to the river and to spawn is basically a political issue involving a negotiation amongst user groups to allocate salmon to their preferred use. However, science-based policy advice about salmon survival in the Columbia River basin surely has to at least acknowledge that those harvests are occurring to have credibility.

In summary, Schaller et al. (and many others in the Columbia) have put more faith in the PIT tag-based "survival" estimates than is warranted. We agree that the PIT tag detectors at the dams probably identify close to 100% of the PIT tagged adults returning to the dams and therefore give a nearly perfect estimate of what *returns*, but the adult return is a badly biased estimate of adult survival because it ignores the large and varying downward bias caused by ignoring harvest that we documented. The uncertainty (variance) surrounding those biased estimates is probably an even worse problem, but the relative degree is currently conjecture.

For reasons perhaps lost to history, the developers of the PIT tag systems chose not to develop the infrastructure necessary to identify PIT tagged fish in the sport and commercial harvest, which would have gone some way towards addressing the problem. As a result, we have to either rely on the CWT-based estimates to estimate the harvest fraction or pretend that they were zero. Thus Schaller et al. are incorrect in their assertion that the PIT tag-based SAR estimates "*are more accurate and precise*" than CWT-based estimates—they might be if, for example, spawning ground escapements are not monitored for CWTs, but this remains to be determined for each population used in our analysis—although we tried to avoid errors like this, it simply wasn't possible to certify that none of the 100+ populations used in the analysis didn't have some issues that we could not identify.

Finally, as a practical matter, although the last quarter century of SAR estimates may be irretrievably lost for use as reliable estimates of survival, this does not mean that this *has* to still apply in future—

the various coastal and river harvests could be surveyed for PIT tags or (more likely) sampled using DNA-based methods. Thus, even if the past estimates cannot be fixed, future estimates can be, at least in principle.

Summary of Methods and Data base problems in Welch et al.

Based on the description of Welch et al. in section 2.2, they calculate the SARs based on the equation in section 2.2. This is a very different estimation procedure than the brood year survivals the PSC calculates, which are the Age 2 survival rates. The age 2 cohort size is estimated through the CWT run reconstruction for a stock specific brood year for a PSC indicator stock. This Age 2 survival rate = (age 2 cohort estimate/CWT release numbers) for an indicator stock and brood year. This is in no way equivalent to SARs used in the Snake or Columbia River studies for stream type Chinook (CSS reports, Petrosky and Schaller 2010, or Haeseker et al. 2012). The Age 2 brood year survival rate was a convention used to estimate a survival factor to be used in the PSC projection model.

So, what Welch presents in the paper is an SAR based on his formula in section 2.2. These SAR estimates are not contained in any PSC documents or produced by the PSC Chinook Technical Committee. Welch carefully states in section 2.2 that the *"database was the source of CWT-based Chinook survival estimates for all regions outside of the Columbia River Basin and for a few stocks located in the Columbia River Basin."* He claims; *the PSC database provides several methods of SAR*. However, review of the supplemental material for the Welch paper provided no reference to a PSC document for his SAR approach. However, curiously in the Welch supplemental material (see attachment Table S1. Datasets of smolt-to-adult return (SAR) estimates for Chinook salmon (*Oncorhynchus tshawytscha*) used in this study: faf12514-sup-0002-tables1.docx) there is a reference to a personal communication with Gayle Brown of DFO (see reference below), **which is not a sanctioned PSC database.** Therefore, they appear to be using something produced by Canada DFO, which is not an official PSC document or database. Either way, it is not clear where the documentation for these data sources and SAR methods resides, or, if they do at all.

KRS: See Appendix 1 for our response to the similar comment by the FPC. See Appendix III for the email trail of our correspondence with the CTC to trace the source of the issue. To briefly summarise, the SARs data that we attribute to the CTC were sent to us by the (now retired) former co-chair of the CTC, Dr Gayle Brown. Although we talked with Gayle multiple times to make sure we understood the data, we were still under the mistaken assumption that they were official estimates from a database. It now appears they were at least calculated using a transparent formula and using data from CTC databases. Additionally, the values we used are closely linear to the CTC's official estimate of the number of immature age 2 (Spring) or age 3 (Fall) ocean age Chinook in the ocean prior to the ocean fisheries occurring. We don't use the CTC estimate of pre-fishery abundance because that measure doesn't estimate the abundance of mature adults and our scientific question is "*What proportion of adult Chinook salmon from the various populations survive from the smolts released?"*, not *"What fraction of the smolts released survive to be in the ocean just prior to becoming vulnerable to the fisheries?"*.

Independent of Welch's misattribution of the source of SAR data and methods, there is a long list of problems with the data Welch uses and the inconsistency for SAR methods he employs. These problems include:

1- The dependence on comparing hatchery stock CWT based SAR performance with wild populations.

KRS: This is false. We never compared hatchery SARs with wild SARs and were careful to keep the comparisons separate throughout the paper.

2- The inappropriate summary and comparison of SAR survival performance using yearling and subyearling populations interchangeably.

KRS: We sequentially and separately discuss the two groups. If Schaller et al. want to identify a place in the text where we didn't do so, we are happy to have the discussion. Otherwise, this claim is false too.

3- - The inconsistent approach for calculating CWT based SARs for Columbia River populations versus the rest of the Oregon, Washington, B.C. and Alaska indicator stocks. Welch states in the paper under section 2.3 that for Columbia River Chinook stocks '*they collated some annual reports.to build up a partial inventory of CWT-based SAR estimates for Chinoo*k." However, by Welch's own description there are big differences from the PSC method and huge discrepancies from hatchery program to hatchery program within the Columbia River basin. Welch states that these supplemental Columbia River CWT based SAR estimates were not expanded for incidental mortality or interdam loss, a departure for the method he applies for estimating SARs for the other Oregon, Washington, B.C. and Alaska indicator stocks. In addition, the supplemental Columbia River hatchery program CWT based SAR estimates that Welch inventories have inconsistent methods applied across these programs. In particular, Welch states in section 2.3 that; 'Hatcheries that do not tag 100% of smolts released may expand their estimates for the proportion tagged while others are estimated using only tagged fish.' See Table S1 for details. Therefore, by Welch's own description, there are very different methodologies applied to estimating CWT-based SARSs across the regions and even within the Columbia River programs. It appears this fact alone would invalidate the conclusions of the range wide comparison on SARs, given that these discrepancies can impact SAR estimates and associated levels of uncertainty.

KRS: Yes, there are differences, which is the reason why we called for a review by the funding agencies (not the fundees) of how SARs are calculated. We were disappointed (and surprised) at the substantial differences in methods that we found in how groups calculate SARs and even in the terminology used. In many cases, the methodologies do not even seem to be documented. (For example, in at least one case we found promising CWT-based "SAR" estimates for a Columbia River hatchery only to discover that the "adult returns" reported were actually only the hatchery rack return number—the number of adults taken for broodstock purposes—and did not include the estimated number of hatchery-produced adults spawning on the spawning grounds). This is unfortunate given the ever-increasing appetite for instituting "SAR" monitoring programs in the Columbia River basin—see our Fig. 8. Similarly, we were equally disappointed to discover just how large harvest levels were and that the CSS often failed to incorporate this.

Some serious soul searching is needed here by the fisheries community because the monitoring programs are costing a substantial amount of public funds but need to be more carefully assessed for rigor and consistency.

However, to respond Schaller et al.'s main point—that there are discrepancies in how SARs are calculated between populations—we agree. However, we did talk with hatchery biologists at almost all Columbia River basin facilities for which we accessed data to ensure that all elements of the return were correctly incorporated. To our knowledge, the only differences between these estimates and those made using the CTC data were incidental mortality (should be small), and inter-dam loss (questionable if it should be included anyway) which are both described in our paper. Whether further data mining will reverse our conclusions is up in the air—this will take more work beyond where we could take the analysis with the available funding.

3 - These previous issues we identify do not touch on the subject of how the CWT-based SARs have a high degree of uncertainty given the problems of high variability in catch and escapement sampling rates that generate extremely wide confidence bounds in the CWT based SAR estimates. Welch does not report any confidence bounds for his CWT based SAR estimates failing to consider the potentially large uncertainty in his methodology. Welch et al. also ignores this uncertainty when comparing SAR estimates across regions and, therefore, draws tenuous conclusions.

KRS: The CTC does not report variance estimates on the survival values. Incorporating these variance estimates could be of interest for a future paper, but will likely be impossible to effectively implement on a coastwide basis owing to the data demands. We believe that the issue of primary policy importance is to look at mean (or median) survival levels along the coast; the variance of those estimates is a secondary issue appropriate for asking how likely is it that the Snake River values are actually lower than other regions despite what the medians indicate.

We did attempt to address the uncertainty in the relative CWT-based estimates in Fig. 4 two ways: (a) by relying on a comparison of median values (geometric means) which tend to be a more robust metric than means for comparison in the presence of outliers, and (b) by using boot-strap resampling of the annual population-specific SAR estimates to create and encompass the underlying uncertainties (Fig. 4). The assumption here is that if the individual CWT survival estimates have substantial variability because of issues with data collection then re-sampling the collection of estimates allows us to better identify how likely it is that some combination of samples lies outside the norm. (In this case, that the bootstrapped samples of regional SARs divided by the Snake River values for the same time period are significantly greater than 1).

The bootstrap technique is a standard approach in modern statistics. Whether it adequately encompasses the true level of uncertainty is obviously up for debate, but the key point of our paper was that the ***median*** SAR values in recent years are mostly similar or below the values reported for the Snake River. That observation, if true, has important policy relevance and is likely less impacted by the degree of underlying uncertainty in the relative SAR values. In other words, no matter what the as-yet unidentified sources of uncertainty are, they aren't likely to reverse the conclusion that hatchery Snake River subyearling Chinook have greater SARs than hatchery Puget Sound stocks, for example, or that hatchery Snake River yearling Chinook SARs are largely similar with other regions although they may change the width of the confidence limits around those regional medians. See our response in this document to the FPC criticism that imbalance in the number of annual SAR observations between regions might somehow result in our finding that SARs are similar in the five year period 2010-2014; there we extend the analysis back to the beginning of the available record, do the comparison of SAR ratios on an annual basis, and find much the same thing.

We hope that the ISAB will address the important sociological issue of why most scientists were not willing to investigate, much less point out, that the numerical values for the SARs for many regions were in many cases worse than the Snake River. (The answer probably lies in not wishing to have to write rebuttals to the FPC's 21 page memo and Schaller et al.'s 16 page memo).

2.2 Pacific Salmon Commission (CWT-based estimates)

Hatchery SARs primarily represent subyearling fall Chinook and are generally calculated from hatchery release to return to hatchery or spawning grounds plus harvested fish. An exception, are five Alaska spring Chinook hatcheries that release directly into ocean after seawater acclimation, which is very different from headwater Columbia River Chinook experience. The wild Alaska stocks are tagged during downstream migration (SARs exclude mortality from upstream natal areas). It would be informative to have the confidence bounds on these stocks (low mark numbers, high uncertainty, no mark/unmark ratio specific to the wild stocks).

KRS: We documented the point that the five Alaskan hatchery stocks are released directly to the sea and for this reason were likely not a good comparison (p. 197: "*Exceptions include five Alaskan hatcheries used in our analysis which are located at sea level and which release smolts directly into the ocean after several weeks of seawater acclimation in holding pens, eliminating losses in freshwater (see later).*") and *"p. 200: The wild yearling Chinook populations in SE Alaska tend to have lower survival than the hatchery-reared population; however, the Alaskan hatchery SAR estimate provided to the PSC is based on combined data for five hatcheries that all release smolts directly into the ocean after acclimation to seawater for several weeks, eliminating losses from freshwater migration").* We do not have data to develop the confidence bounds Schaller et al. mention, but it could make for an interesting follow-on paper. The data are presumably available from ADFG.

2.3 Agency estimates (CWT-based estimates)

Welch et al. used published estimates for fall, fall-winter, winter Chinook from Sacramento River. They collated hatchery reports to estimate SARs for Columbia Basin spring Chinook (SARs exclude upstream passage mortality). Questionable how well these methods align with Welch methods employed for PSC indicator stocks and quality control on CWT recovery methods. Methods and quality control for many of the supplemental groups differ dramatically. Some of the CWT based estimates were for experimental groups, which is inconsistent with PSC indicator stock estimates for production groups.

KRS: "Agency" estimates were used after discussion with hatchery biologists to ensure the data were representative and that all elements of the run were incorporated and properly expanded. Names of the hatchery biologists are included in the Acknowledgements. We were not made aware of any experimental manipulations that the hatchery biologists felt would bias their results.

We agree with Schaller et al. that having more consistent methods across datasets would be good thing. However, the relative survival values we report have been going on for years while more and more data was collected. No one seems to have thought it worthy to point out that the relative SAR levels don't fit with expectation. It was for this reason in our paper that we call on the funders to conduct a review and ask what has gone wrong. There seem to be two choices here: (a) The data isn't very good, raising the question of why 2,279 years of effort may have been wasted or (more likely in our view) (b) no one was prepared to point out that the collected data did not fit with some long-standing beliefs.

Given the importance of the implications stemming from the numbers, we agree that it would be a good thing to look further into how all these monitoring programs are being conducted. We even explicitly called for that in our paper.

2.4 <u>Pacific States Marine Fisheries Commission Estimates</u>

The authors used estimates for UCOL Entiat hatchery spring Chinook. It is unclear what data quality control was employed. The PSMFC source is a data base, and the authors don't describe how they calculated SARs. Entiat mark numbers were variable over years, and in recent years CWT marks were significantly reduced (no CWT estimates after brood year 2007). Entiat Hatchery moved to PIT tags to better estimate survival rates for the hatchery. It appears the Entiat CWT expansions done by authors had no explanation of approach, no Confidence Intervals provide for CWT based SARs, and no mark levels reported by authors. Entiat CWT marking ended in 2007 (Table S1), why is this reported then in 5 most recent years (2010-2014)? The authors' approach for estimating Entiat SARs appears to be a big outlier from the approach they employed to estimate the SARs for PSC indicator stocks.

KRS: As we state in the manuscript (Section 2.4), we used the PSMFC's RMIS database "*…only for Entiat Spring Chinook after consultation with Entiat Hatchery biologist on the integrity of the dataset*". We called Entiat Hatchery to ask if there was any reason the SARs calculated directly from the RMIS database might be biased (e.g. from use of experimental releases) and if all components of the run were included and were properly expanded for the sampling proportion. In retrospect, we might have been clearer on the specific method of data access from RMIS and subsequent calculation of the SAR ratio. At the time, it seemed straightforward and similar to methods we had described for the PSC and Agency estimates (Section 2.2 and 2.3): download and total the releases, download and total the expanded recoveries, and calculate the ratio. (Note that the recoveries are already expanded for the sampling effort in the RMIS database). Confidence intervals are not provided.

Entiat CWT marking for spring Chinook were excluded from the five most recent years (2010-2014) because the program ended in 2007. We don't know why Schaller et al. think they were included in this particular dataset.

2.5 <u>Raymond (1988) estimates</u>

Used estimates for Snake and upper Columbia stream-type Chinook (1960s-80s). Smolts indexed at uppermost dam, adults as returns to uppermost dam plus harvest loss (excludes upstream passage mortality as well as mortality of juveniles from natal areas to uppermost dam). Note, Raymond's SAR estimates for Snake stream-type Chinook have been re-calculated to also include upstream passage mortality in CSS reports, but these re-calculated estimates were not used by Welch et al. 2020.

They claim they use the Raymond data in conjunction with CWT estimates for a more complete time series. CWT's don't begin until 1979 when comprehensive PSC recovery began.

KRS: We didn't mean to imply that both the Raymond and CWT estimates were available at the same point in time. We included the Raymond data for a more complete (i.e. longer) time series.

We were unaware that the Raymond estimates had been recalculated to include upstream passage mortality. The Raymond estimates we used estimated adult returns fairly far up the river: to Priest Rapids Dam for the Upper Columbia stocks, and to Ice Harbor Dam for the Snake River stocks.

Note also SARs in upper Columbia and Snake had declined by the early 1970s in response to dam construction and operation, this decline was prior to the 1978 ocean regime shift.

KRS: As per Figure 2, there was some decline still occurring around the 1978 regime shift. We refer to the 1978 regime shift not as the cause of the decline, but as a reference point to classify "longer" time series. Construction of the dams likely reduced survivals, but these effects seem to be mitigated. SARS are currently fairly similar for most regions of the coast such that regional factors are not likely the cause of current survival levels.

2.6 Comparative Survival Study (PIT tag-based estimates)

Welch cites using SARs for Snake river Chinook estimated from Lower Granite back to Lower Granite. Note that SARs are also available from LGR to BON, which for wild stream-type Chinook represents primarily pre-harvest returns since the 1990s. (Since ESA listing, there is no directed harvest and limited catch and release mortality (about 2%) allowed below Bonneville dam).

KRS: Adult returns to Bonneville Dam are less representative of survival over the migratory cycle than LGR to LGR. (As the ISAB will know from our response to the FPC memo, several attempts to obtain more representative data from the FPC that included the above dam smolt survival component were stonewalled). Using LGR-BON survival estimates increases the dissimilarity from survival measured using other methods, so we didn't use it—as much as possible we wanted to minimize the potential criticism that the data weren't comparable.

Welch et al. cite CSS report language that SARs do not incorporate losses due to harvest because PIT tags not sampled in fisheries. They neglect to mention or cite that several published analyses (and Chapter 5) do incorporate harvest rates (and upstream passage loss) in stream-type Chinook and steelhead SAR estimates (a fairly trivial matter for stream-type Chinook and steelhead).

KRS: While the published papers may account for harvest, they are applicable only Snake River spring Chinook. As noted above, the CSS produces numerous SAR estimates. Further, Schaller et al.'s claim that the matter is trivial is not true, even for stream-type Chinook and steelhead. We found that Snake River wild Spring Chinook had the lowest reported harvest rates of any group. However, these harvest rates varied over time from ~10% to ~20% since 1998, resulting in the need to multiply SAR estimate by ~1.1X to ~1.25X to compensate, as we pointed out in the paper. For other populations the impact is much larger. These are substantial corrections compared to the tiny changes in adult survival expected from spill or dam modifications affecting the smolts.

2.7 Division by life history

Populations were grouped by subyearling/fall and yearling spring Chinook. When Welch et al. draw the conclusion that ocean conditions are primarily responsible for the patterns of Chinook survival coastwide, it is inappropriate to lump their SAR estimates for stream-type and ocean type Chinook stocks. In addition, they ignore the numerous studies that assess the ocean and inriver factors that best explain the variation in SARs and life cycle survival for wild stream type Chinook.

KRS: This criticism is false. As we noted in our response to a similar claim in the FPC memo, we devoted an entire section of the paper (§2.7) to discussing why we need to treat the two groups separately. As for the claim that we "*ignore the numerous studies*", we cited some of them, but all of these are statistical exercises based on either correlation or regression analysis (which is essentially the same thing—one mathematical way to calculate the correlation is as the geometric average of the slopes of the regressions of Y on X and of X on Y, $r = \pm\sqrt{b_{XY} \cdot b_{YX}}$, where b is the slope of the respective regression[3]). We can agree with the authors that they have identified interesting statistical associations amongst the variables, but correlation is not proof of causation. To achieve the latter will require conducting explicit hypothesis tests rather than correlations.

2.8 Comparisons between regions

The authors do not document how they generated consistent estimates for SARs within and across regions. In fact, the authors conclude that 'we encountered substantial challenges in fully understanding whether all components of adult returns were adequately included in many SAR time series. This appears to be a fundamental flaw in the coastwide CWT based SAR analysis. This problem only compounds when the authors ignore the uncertainty for these CWT based SARs, then they attempt to evaluate patterns of survival across regions, and then compare and make judgement about the efficacy of much more precise PIT based SARs. The authors make no attempt at formally evaluating the similarity or difference in patterns of SAR estimates across populations. In fact, no attempt is made to estimate the CIs for any of the SAR estimates. The authors simply draw the conclusions for their interpretation of a visual inspection of figure 2. The conclusions drawn from similarities and differences from figure 3 do not provide an evaluation of the pattern of SARs. Also, given the differences in life histories, hatchery vs wild, and within hatchery groups experimental versus production groups the SAR comparisons are dubious at best.

KRS: Patently false. The claim here is that "*The authors make no attempt at formally evaluating the similarity or difference in patterns of SAR estimates across populations. In fact, no attempt is made to estimate the CIs for any of the SAR estimates.*" Actually, we used bootstrapping to derive the confidence intervals around the ratio of SARs between regions in the five year period 2010-2014 (Fig. 4 and Supplementary Info Fig. S1), which addresses the key issue—how similar SARs are from different regions? Fig. 2 shows the time series for the readers, so that they can visually look at the data. Fig. 3 is the box and whisker plot comparing the median CWT based-SARs and the distribution of the data

[3] See, for example, Chapter 300-9 of https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjSmK-nuqvuAhWjFzQIHV5zBp0QFjARegQIMxAC&url=https%3A%2F%2Fncss-wpengine.netdna-ssl.com%2Fwp-content%2Fthemes%2Fncss%2Fpdf%2FProcedures%2FNCSS%2FLinear_Regression_and_Correlation.pdf&usg=AOvVaw2zM294eKi_fjC6dCpO-yFt

around the medians. Fig. 4 then shows the statistical test of the SAR ratios. We arranged the analysis this way because it is good practice to look at the data before blindly calculating confidence intervals… The confidence intervals that Schaller et al. claim we didn't do are explicitly shown in Fig. 4.

We agree with Schaller et al.'s point that the uncertainty in the CWT-based SAR estimates is complex. One approach would have been to go into the underlying data and try to develop clear formulae for the uncertainties. But given that CWTs have been used for over half a century and this is not routinely done is probably sufficient warning that it cannot be successfully done without a plethora of assumptions. Instead, we opted to use resampling approaches to ask what the uncertainty was in the observed SARs. If the individual annual SAR estimates are inordinately variable, then the resampling approach should generate empirical 95% confidence intervals that are "unreasonably" wide compared to what we measured. First, they didn't. Second, the key policy issue is the relative survival levels amongst regions relative to the Snake River region. Schaller et al. are essentially hoping that the uncertainty is so large that the possibility of Snake River SARs being below the lower 95% confidence interval is still plausible.

Some perspective may be useful here. Schaller et al. (and the FPC) are arguing that because the data contain complexities "maybe" our results aren't real. Fair enough. But the basic data are pretty simple—how many adult salmon survived to adulthood divided by how many smolts went to sea? In order to generate consistent estimates, we included only PSC indicator stocks supplemented by published research and with agency estimates verified by personal communications with hatchery biologists. We think that the most serious problem is that although that ratio is now similar coastwide, no one was willing to make that statement before now, precisely because of the furor that pointing out the obvious generates. Arguing that the data are "complicated" doesn't explain why SARs have converged over time to about the same level.

The extremely short 5-year time series (employed by the authors) inhibits making robust and rigorous conclusions about ocean factors that influence survival patterns for Chinook populations. Because of different mark levels and recovery rates for different groups, the analysis should have directly considered the CI about the CWT SARs to reflect the highly uncertain nature of these estimates.

KRS: We chose the most recent five year period because we wanted to specifically show recent SAR levels across regions. In our response to the FPC (Appendix I), we added a new figure that shows the annual SAR ratio relative to the Snake River for individual years extending back for all available years on a region-by-region basis; our conclusions remain the same. The bootstrapping method we used, calculated across all the data, just do not show that results that will fit with what Schaller et al. believe.

The use of normalized ratio relative to Snake isn't clear whether this is for CWT. Analyses should directly incorporate the uncertainty in CWT and PIT SARs, to support rigorous and robust conclusions concerning temporal and spatial survival patterns for Chinook populations.

KRS: The first sentence in the legend to Fig. 4 helpfully states "*Regional CWT-based SAR estimates for Chinook salmon normalized relative to Snake River SARs for the 2010–2014 period*" The y-axis was similarly labelled as "*Normalized CWT SAR*".

As we have already pointed out, we used a bootstrap resampling method to incorporate the uncertainty in the CWT and PIT SARs. In the case of Fig. 4 the legend further states "*Horizontal red lines show the empirical 5% and 95% percentiles on the sampling distribution of the normalized ratio*".

Note: The authors acknowledged that SARs may not measure the same portion of life cycle for the various stocks employed in their comparison.

KRS: Indeed we did. Thanks. To remind the reader of what the issue is, we pointed out on p. 209 (paraphrasing): "*There are two reasons for this. First, for dam-to-dam estimates the survival losses incurred upstream of the dam can vary substantially between populations. Unless census points are located at the start and end of the migration period, the amount of excluded upstream survival acts as a population-specific random variable influenced by the excluded distance" and "The second reason is that Chinook harvested in fisheries prior to return are not accounted for in PIT tag-based estimates". From at least one perspective, both can be viewed as limitations imposed by the choice to use PIT tags (although above-dam smolt survival could readily be included in survival comparisons if desired). We view the deficiencies as being introduced by the choice to use PIT tags, although we acknowledge that CWTs have significant issues as well. In our view the matter comes down to one of focus—if survival for the migratory phase is desired, CWTs are probably preferable at present. If focus is restricted to smolt or adult survival while migrating within the hydropower system, then PIT tags are clearly preferable.*

2.9 Comparison between CWT and PIT tag-based SARs

It is not clear from Welch et al. how they made adjustments for release to LGR and harvest mortality PIT based SARs.

KRS: We made no adjustments. The regressions are for the unadjusted data (note the very high R2 values). As we state in our response to the FPC memo, we were stonewalled in attempts to incorporate adjustments in smolt survival from release to LGR. Also, the harvest rate data we found were so limited that we did not want to attempt to make partial corrections.

In the equation for the aggregate correction factor in section 2.9 it is unclear and undocumented where the estimates for the parameters S_{smolt} and S_{adult} originated, which are critical parameters that shape comparison.

KRS: Apparently, this section of the text was not worded clearly enough. These correction factors weren't calculated (in large part because the FPC stonewalled attempts to determine smolt survival upstream of the dams). What we are doing in the stated equation in section 2.9 is describing the three factors missing from PIT tag-based SARs. These missing components should be captured by CWT-based survival estimates. The slopes of the 12 separate regressions shown in Fig. 6 quantify those combined values for each population. The high R2 values demonstrate that the combination of the three missing factors from PIT tag-based SAR estimates is probably sufficient to closely match the CWT based estimates.

The second approach to estimate correction factor was to identify populations with both CWT based and PIT based SARs through regression analysis. The authors tried, but couldn't develop a relationship between PIT and CWT based SARs. They couldn't find much data from the same populations to compare CWT based to PIT based SARs, and all of these comparisons are from hatcheries. Where both SAR estimates were available, regression relationships were strong but biased. Subyearling CWT SARs were higher than subyearling PIT SARs (because harvest was not captured in PIT SARs). Yearling CWT SARs were lower than PIT SARs. A correction factor appears infeasible. All of these comparisons lack rigor, because the authors ignored the confidence intervals for CWT based SARs vs CI for PIT tagged based SARs. Ignoring uncertainty for CWT based SARs (which tends to be relatively high) vs PIT based SARS (relatively low) and the small amount of comparison data, brings the authors slim findings into question.

KRS: "*The authors tried, but couldn't develop a relationship between PIT and CWT based SARs*". We disagree—we found an excellent one, but it was population-specific. As Fig. 6 shows, all but one R^2 values are ≥88% (one of 12 populations had an R^2 of "only" 82%). This is excellent agreement and demonstrates two things: (a) that the differences between PIT & CWT-based SAR estimates are consistent for a given population, and (b) that the deficit is in the PIT tag-based estimates… as we stated in our paper (p. 202): "*…PIT-based estimates differ in two major ways from CWT estimates: (a) they exclude sport, commercial, and indigenous harvest and (b) they exclude smolt and adult losses in the region lying between the uppermost dam and the hatchery or spawning site.*". So there are excellent, simple relationships evident. What we could not develop was a "*one size fits all*" relationship because of the factors that the PIT-based SAR estimates exclude are population-specific. That is not at all the same thing as what Schaller et al. are stating.

3 Results

They used SARs from 94 hatchery populations, 26 wild and 3 hatchery-wild mixed populations. All populations outside Columbia are CWT-based. Within Columbia, CWT and PIT SARs were used.

3.1 SARs from coded wire tags

Independent of Welch's misattribution of the source of SAR data and methods, there is a long list of problems with the data Welch uses and the inconsistency for SAR methods he employs. These problems include:

1- The dependence on comparing hatchery stock CWT based SAR performance with wild populations.

KRS: False. We never compared hatchery populations with wild populations. There is simply nowhere in the paper where that is done.

2- The inappropriate summary and comparison of SAR survival performance using yearling and subyearling populations interchangeably.

KRS: We don't mix up the two life history types, despite what Schaller et al claim. See section 2.7 for an explicit discussion of the reasons why we kept them separate.

3 - The inconsistent approach for calculating CWT based SARs for Columbia River populations versus the rest of the Oregon, Washington, B.C. and Alaska indicator stocks. Welch states in the paper under section 2.3 that for Columbia River Chinook stocks *'they collated some annual reports......to build up a partial inventory of CWT-based SAR estimates for Chinook."* However, by Welch's own description there are big differences from the PSC method and huge discrepancies from hatchery program to hatchery program within the Columbia River basin. Welch states that these supplemental Columbia River CWT based SAR estimates were not expanded for incidental mortality or inter-dam loss, a departure for the method he applies for estimating SARs for the other Oregon, Washington, B.C. and Alaska indicator stocks. In addition, the supplemental Columbia River hatchery program CWT based SAR estimates that Welch inventories have inconsistent methods applied across these programs. In particular, Welch states in section 2.3 that; *'Hatcheries that do not tag 100% of smolts released may expand their estimates for the proportion tagged while others are estimated using only tagged fish.' See Table S1 for details.* Therefore, by Welch's own description, there are very different methodologies applied to estimating CWT-based SARSs across the regions and even within the Columbia River programs. It appears this fact alone would invalidate the conclusions of the range wide comparison on SARs, given that these discrepancies can impact SAR estimates and associated levels of uncertainty.

KRS: Because methodologies vary for estimating the ratio of adult returns to smolt releases does not *"…invalidate the conclusions of a range wide comparison*". We used the data that we could reasonably find after verifying that there were no major biases. If the CWT data were as bad as Schaller et al. claim, we would not have obtained such high R^2 values in the regression of CWT-based SAR estimates on Schaller et al.'s "*highly precise"* PIT tag-based SAR estimates. So, *ipso facto*, the CWT data appear to be fairly good for comparing between regions despite their warts. (And, to drive the point home for the readers that have gotten this far, the really big deficiencies we found are in the PIT tag-based SAR estimates: failure to account for harvest and exclusion of variable amounts of the life history in the reach between the hatchery and the top-most dams.

As for the claim that "supplemental Columbia River CWT based SAR estimates *were not expanded for incidental mortality or inter-dam loss, a departure for the method he applies for estimating SARs for the other Oregon, Washington, B.C. and Alaska indicator stocks*", in our estimation the effect will be trivial within the comparisons and should not materially affect the conclusions. As with all other aspects of the paper, we have provided the raw data. Schaller et al. are welcome to demonstrate that their concern has substance.

4 - These previous issues we identify do not touch on the subject of how the CWT-based SARs have a high degree of uncertainty given the problems of high variability in catch and escapement sampling rates that generate extremely wide confidence bounds in the CWT based SAR estimates. Welch does not report any confidence bounds for his CWT based SAR estimates failing to consider the potentially large uncertainty in his methodology. Welch et al. also ignores this uncertainty when comparing SAR estimates across regions and therefore draws tenuous conclusions.

KRS: See the prior responses. Confidence bounds were calculated in the most robust way we know how to do: bootstrap resampling. In any case, this issue of the inherent uncertainty in the SAR estimates is a red herring. Because regional SARs for most other regions were generally lower than the Snake River (subyearlings) or similar (yearling) then expanding the width of the confidence intervals will just make the possibility of demonstrating any differences in SARs that actual supports what Schaller et al. believe even more statistically remote. What is needed is a demonstration of a reason why over time measures of central tendency (mean/median) for the SARs from most other regions have fallen from higher levels to approximate the Snake River values. So far we can't identify one—and neither have Schaller et al.

They state that SARs extending back to before 1978 show 3-fold decrease in SARs for hatchery populations. Snake and Upper Columbia rivers stream-type Chinook were in serious decline by early 1970s, well before 1978 regime shift]. Not clear what baseline they actually used for this. Selection of a base line needs to be clear so as to precisely test a posited hypothesis. Methods and results are organized around a vague set of questions without a specific approach to evaluate a well-defined question or hypothesis. For example, the authors describe how they '*collate Chinook SAR time series for the west coast of North America to document broad patterns in survival*'.

KRS: The key hypothesis we tested was whether SAR levels were closely similar between regions and particularly with respect to Snake River survival, as the abstract hints at: "*…the SARs of Snake River populations, often singled out as exemplars of poor survival, are unexceptional*". In the context of this hypothesis test, a SAR ratio significantly greater than one (equality) would have been evidence that other regions without dams have better survival. However, as the abstract went on to state: " *[the SARs of Snake River populations are] in fact higher than estimates reported from many other regions of the west coast lacking dams*". And as for the secondary issue (the 3-fold magnitude of the decline on the longer time series), this is readily apparent from Fig. 2 for longer time series (i.e., those that extend back to the 1970s). SARs are plotted in Fig. 2 using a $\log_4$ axis, so a 3-fold change in SARs is $\log_4(3)\approx0.8$ of one unit (space between tic marks).

Wild populations have higher SARs than hatchery. Limited CWT data for wild and no data from wild versus hatchery SARs from same population. Wild AK SARs lower than hatchery SARs but AK hatcheries release directly into ocean. Only CWT populations in 2-6% recovery range was UW experimental hatchery in Puget Sound & Chilliwack hatchery in Strait of Georgia.

KRS: This is a summary of some of the results from our paper. No response necessary.

<u>3.2 Comparison between regions</u>

Authors conclusion that ocean conditions drive these coastwide SAR patterns and that freshwater conditions may not be drivers of SARs, is questionable given all the methodological flaws described above; In addition, these conclusions are reached by the authors visual observation of patterns in their Figure 2. In order to support their conclusion a rigorous statistical analysis of these survival rate times series is warranted;

KRS: Perhaps Schaller et al. should read the text around Figs. 4 and 6 (last time we looked, bootstrap (Fig. 4) and regression (Fig. 6) both counted as "rigorous statistical analysis"). And, as we point out above, the claimed methodological flaws didn't prevent us from finding R2 values of ≥88% in 11 of 12 populations (82% for one more). For field data, these are remarkably high values. Schaller et al. are silent on (a) why such high consistency between PIT & CWT tagging methods occurs, and (b) why the departure from a 1:1 relationship are chiefly attributable to deficiencies in the PIT tag methodology (failure to account for harvest or above dame smolt & adult survival).

Authors (dubiously supported) conclusion that ocean conditions drive these coastwide SAR patterns, is contrary to the large body of published literature that concludes Chinook salmon population survival rates are most influenced during estuary and early ocean life stages (numerous references from our past pubs; Particularly Petrosky and Schaller 2010). Recruitment success in the ocean environment is generally believed to occur largely during the first critical months at sea (Ricker 1976; Nickelson 1986; Pearcy 1992; Mueter et al. 2002, 2005; Pyper et al.2005; Peterson et al. 2006).This early marine stage is a very vulnerable life stage for salmon because they make the transition from freshwater phase to a seawater phase involving numerous physiological changes while encountering marine predators. Therefore, especially recognizing the various dispersed ocean entry points for the salmon populations included in the Welch et al. paper, the influence of estuary and near shore ocean conditions would be highly variable (Peterman et al. 1998 shows strongest correlation of marine survival within about 500 km of ocean entry points), and would not support a common marine influence across such a wide geographic range of populations used in the Welch et al. paper.

KRS: It is useful to recall that "*survival rates are most influenced during estuary and early ocean life stages*" is *an assumption*. The hypothesis was put forward by the eminent Norwegian biologist Johann Hjort (Hjort, 1914). Testing Hjort's critical period theory requires (a) measuring smolt survival just after the end of the *"estuary and early ocean life stage",* and (b) showing that it is highly correlated with survival at adult return. With the exception of Kintama's earlier POST telemetry studies, there is literally no data on the planet to back up the theory one way or the other

The point of our response here is simply a reminder that almost all critical period studies, much like the many papers Schaller et al. cite above, are correlation exercises and do not demonstrate cause and effect. (Incidentally, it is well-recognized that scientific studies not fitting with pre-expectation often get hostile reviews not levelled at statistical exercises fitting expectation. The dearth of papers not finding a statistical correlation supporting the critical period theory may just reflect the difficulty in getting negative results published).

The interested reader should examine the biomedical literature where there is now a concerted effort to move away from precisely the sort of correlation-based analyses that Schaller et al. cite. We think that the stakes are too high in the Columbia River to manage the fate of the hydropower system on the basis of statistical correlation and unvalidated assumptions, such as the one that the early marine phase is "critical". There is, for example, accumulating evidence that the dramatic increase in pinniped predation over the past few decades can potentially cause a great deal of mortality on returning adults. If all regions of the coast had similar increases in adult pinnipeds congregating in the lower reaches of rivers to intercept returning adults, this could be one way to achieve a pattern of decline in survival to similar levels in recent years. (We are not defending this as the primary mechanism driving the poor returns, merely using it as an illustrative of how the "salmon problem" does not have to be caused by a "critical period" in the early marine phase).

However, the only comparison tests the authors' implement is using a very short time series (recent 5 years) without taking into consideration the high level of uncertainty for CWT-based SARs. It is difficult to see how from a short 5-year time series analysis (that ignores the confidence intervals about SAR estimates) that one could seriously evaluate the influence of ocean conditions on west coast Chinook survival rate patterns. Finally, the authors claim 'they examined the CWT and PIT tag SAR data sets to evaluate the broader evidence for "delayed mortality," an important theory that argues that the greater dam passage experienced by Snake River stocks predisposes these populations to lower subsequent survival after migration out of the hydropower system than populations not migrating through the Snake River dams'. The authors present no formal evaluation or hypothesis test for delayed mortality and they dismiss the findings of numerous peer reviewed studies that found the operation and configuration of Columbia River hydrosystem impacts SARs and life cycle survival (delayed hydrosystem mortality) for Snake River stream-type Chinook populations. The results and conclusions of the Welch et al. paper are reached by the authors without any rigorous analyses of SAR estimates or synthesis of previous studies.

KRS: Schaller et al. are repeating the same claims they made earlier in this memo. See our earlier rebuttal of the claim that we did not do a "formal analysis" or "hypothesis test".

SARs from PIT tags

The authors summarize PIT-based SARs in the Columbia River stating that wild fish have generally higher survival and different regions have similar of lower SARs to Snake River. Exceptions are two Mid-Columbia River wild yearling populations (John Day & Yakima), which fall within 2-6% SAR target.

The authors state that both wild and hatchery subyearling SARs from the Mid-Columbia have SARs that fall well below Snake River SAR medians. In addition, the authors state that all other populations (including 3 hatchery Mid-Columbia yearling populaitons) have SARs which rarely or never exceed 2%.

Authors mischaracterize NPCC 2-6% SAR goals. The goals were established for stream-type wild Chinook, and also applied to wild steelhead. It's widely recognized in the Columbia Basin that the specific goal may not apply to subyearling Chinook because of the dramatically different life history strategy and is certainly not a goal to be applied to hatchery populations.

KRS: We simply show the 2-6% goals as bands on the SAR plots because it helps the reader assess what population groups might reach these targets (not many, obviously). In our view, it is just as reasonable to show those on the subyearling comparisons as it is to show them on the Spring Chinook survival values for other regions. For example, the simple fact that we plotted the 2-6% bands on the Spring Chinook SAR comparisons for say, Puget Sound SARs, does not mean that we are suggesting that these recovery goals should be applied to Puget Sound populations. Rather, we chose this presentation so that the reader could evaluate for themselves whether any other regions or population groups were achieving these ambitious goals. They are not, so it may be time to re-evaluate their current use in the Columbia.

3.3 Comparison of CWT and PIT-based SARs

The authors tried, but couldn't develop a relationship between PIT and CWT based SARs. They couldn't find much data from the same populations to compare CWT based to PIT based SARs, and all of these comparisons are from hatcheries. Where both SAR estimates were available, regression relationships were strong but biased. Subyearling CWT SARs were higher than subyearling PIT SARs (because harvest not captured in PIT SARs). Yearling CWT SARs were lower than PIT SARs. Correction factor appears infeasible. All of these comparisons lack rigor, because the authors ignored the confidence intervals for CWT based SARs vs CI for PIT tagged based SARs. Ignoring uncertainty for CWT based SARs (which tends to be relatively high) vs PIT based SARS (relatively low) and lack the small amount of comparison data, brings the authors slim findings into question.

KRS: Schaller et al. are repeating the same claims made earlier in this memo. See our earlier response. We find it remarkable that they repeat the assertion that we "couldn't develop a relationship" when our results using simple linear regressions had R^2 values almost always exceeding 88%.

DISCUSSION

4.1 SAR comparison

3rd paragraph: "North American decreases in survival have occurred despite governments' best attempts through harvest regulation, hatchery enhancement and habitat restoration. A major assumption is that freshwater habitat degradation … make important contributions to the decreasing survival…" Welch et al. ignore the literature in Snake and Columbia that indicate both freshwater and ocean conditions are important drivers. They ignore the NPCC strategy of freshwater actions to ensure salmon can survive in face of variable ocean conditions.

KRS: We cited some of that literature, which is all based on statistical correlations between measures of freshwater & ocean indices and adult return rates (SARs). We have no disagreement that both ocean and freshwater conditions are important get adults back. Our interest is in the major drivers of current poor SARs. As we describe in a previous publication (Welch et al, 2011), events happening in the later ocean phase are ~27 times more influential on SAR than those in the freshwater and early marine phase.

In the last decade, a revolution has occurred in biomedical research that has now recognized that the correlation-based statistical methods research papers have used are subject to far too many problems (subconscious "cherry picking" of data, "p-hacking", and other sins). See (Horton, 2015) for a thoughtful, if despairing, commentary by the Editor of "The Lancet", which we quote in part here: "*The apparent endemicity of bad research behaviour is alarming. In their quest for telling a compelling story, scientists too often sculpt data to fit their preferred theory of the world. Or they retrofit hypotheses to fit their data.*").

8th paragraph: Delayed mortality paragraph: Welch et al. do cite the CSS chapter on delayed mortality & emphasize that direct tests of the theory have not found evidence to support it (citing their own [flawed] studies. They neglect to cite Haeseker's (2013) review comments and also neglect to mention the weight of evidence that actually does support delayed mortality. They further treat the John Day and Yakima wild yearling Chinook SARs as outliers, by stating that three PIT tagged hatchery mid-Columbia yearling populations and two (wild) upper Columbia populations have similar SARs to the Snake River populations (Figure 5). The flaw with this logic is two-fold: (1) hatchery yearling Chinook SARs are typically lower (sometimes dramatically) than wild due to a number of hatchery operational conditions and genetic differences, which can reduce fitness, and (2) upper Columbia wild yearling Chinook pass through similar number of dams as Snake River populations (the similarity in SARs would be expected due to presence of similar numbers of dams). In addition, mid-Columbia wild steelhead also exhibit higher SARs than counterparts in the Snake and upper Columbia (CSS 2019), mirroring the situation for wild yearling Chinook.

KRS: We responded (Rechisky, Welch, & Porter, 2013), to Haeseker's (2013) review of our "flawed" studies (Rechisky et al. 2009, 2013) and then followed them with another paper (Rechisky, Welch, Porter, Hess, & Narum, 2014) that addressed many of Haeseker's stated concerns but that reached the same conclusion (no evidence of reduced Snake River smolt survival due to extra dam passage).

In our response to Haeseker (Rechisky et al., 2013), we pointed out that our PNAS paper was a formal experimental test of treatment and control groups (Snake vs Yakima R smolts) and found no difference in survival to as far away as the northern tip of Vancouver Island, some 1,500 kms away and about 6 weeks after release. That study used carefully size-matched and timed release groups. Haeseker's criticism was that it wasn't actually the difference in the number of dams that caused the difference in survival, it was some subtle ecological difference between the smolts from the two river systems (perhaps size) that when combined with the difference in the number of dam passages precipitated the 3-fold difference in survival.

So here we can see the value of a rigorous experiment instead of correlation-based science-- the argument shifted to one of "*We are still right, it's just an unidentified ecological difference interacting with the amount of dam passage to cause the reduced survival of Snake River salmon*".

Forgive our cynicism here, but if regional biologists seriously thought that Haeseker's argument had merit they would have tried to identify what that mysterious ingredient was. But perhaps they were simply discouraged by our subsequent paper (Rechisky et al., 2014) where we tagged essentially all sizes of smolts simultaneously arriving at a dam and compared post-release survival using genetic techniques to identify how the Snake River smolts fared relative to the other groups. Again we found no evidence of reduced Snake River smolt survival due to extra dam passage, with fewer of Haeseker's stated concerns applicable. Oddly, that paper is studiously ignored in the Columbia, despite addressing many of the misgivings expressed by Haeseker towards using rigorous experimental tests.

In short, if the region still really thinks that delayed mortality is caused by the Snake River dams they should move to identify what the ecological difference was that combined with dam passage to cause the poor survival. Biologists instead continue to believe in an interesting theory. We think that by doing so they may be unwittingly dimming the prospects for real salmon conservation as ocean conditions worsen.

4.2.2 PIT tag-based estimates

Welch et al. statement about exclusion of survival from release to Lower Granite Dam can vary substantially cited (Faulkner et al. 2017). Note that SARs from upper dam to Columbia River mouth (including harvest & dam passage mortality) explained the majority (80%) of variation in life cycle survival rates of Snake River stream-type Chinook over a 70-year period (Petrosky et al. 2020). Whatever survival variation from rearing areas to the uppermost dam occurred in these years was much less influential.

Welch et al. statement that PIT-based SARs do not provide credible measure of SAR (because harvest is not included) is flat-out wrong. SARs can be measured at various locations, as is clearly pointed out in the CSS report which they cite. For wild yearling Chinook, the SAR with returns to Bonneville Dam include virtually all harvest impacts. Since ESA listing, no direct harvest occurs downstream of Bonneville Dam, and incidental catch and release mortality has been capped at about 2% (depending on projected run size). Moreover, incorporating the effects of stream-type Chinook harvest (and upstream passage mortality) into SAR analyses is straightforward and common in CSS Reports (Chapter 5) and peer- reviewed literature (e.g., Schaller et al. 2014; Petrosky et al. 2020).

KRS: Our point is that PIT-based SARs may not provide credible measures of survival, but rather adult returns from the fisheries (escapement). The distinction is important. We agree that PIT tag detectors at the dams likely provide complete detection of adults returning with PIT tags. However, the fisheries are harvesting a large and variable proportion of salmon that would otherwise survive to return. There are a wide range of populations whose SARs that CSS report, not just the Snake River Spring run. Those losses are unaccounted for because of the failure to design a PIT tag monitoring system for sport and commercial catches back when the PIT tag system was rolled out.

To repeat one of the main points of our paper, because the harvest levels we document are much larger than we had been led to believe, statistical analyses of SAR fluctuations actually measure the combination of dam and ocean-induced fluctuations in survival as well as the fluctuations induced by harvest. This on its own is troubling. However, the terms of the Pacific Salmon Treaty introduce a positive feedback system on those harvests that dampen (at best) or reverse (at worst) manipulations of the hydrosystem that might actually affect survival to adulthood for many of these populations.

Outside of salmon fisheries, we know of no fisheries where responsible adults would accept that analyzing the output from such a system was reliable. In short, the PIT tag situation in the Columbia, as we reported it in the paper (and expanded on this response) is likely not credible as currently used for most populations whose SARs are used as measures of survival. This may include the Snake River Spring Chinook, which we agree had the lowest harvest rates of those populations with harvest rate data that we could find. But the fluctuations in these harvest rates and their impact on SARs, even if "smaller" in absolute terms, may still be substantially larger than the direct effect of such dam manipulations as are currently underway, which are based on projection of SARs (not survival) under spill to levels far outside of previous experience.

4.3 Harvest and PIT-based SARs

Welch et al. statement about PIT-based SARs not accounting for stream-type Chinook river harvest is misleading (see previous comments). Harvest accounting for subyearling Chinook PIT-based SARs is more challenging than for yearlings. The authors apparently do not understand how PIT based SARs are used in the Columbia Basin, but uses (or misuses) them anyway.

The authors state that a challenge with using PIT-based SAR estimates to set quantitative recovery targets for Columbia Basin Chinook is that the fishery management strategy is divorced from the goals. This is false for stream-type Chinook. CSS analyses (Chapter 5) and Petrosky et al. (2020) demonstrated that pre-harvest SARs of 4% (the NPCC average SAR goal) for Snake River stream-type Chinook would result in about 70% of historical productivity. In addition, Lower Granite to Lower Granite SARs of less than 1% (after river harvest) have been shown to result in generational declines in abundance of Snake River wild stream-type Chinook and steelhead (CSS Chapter 5).

KRS: We can agree with Schaller et al. that the "pre-harvest" SARs of about 4% would be nice to return to. Where we differ is on the question of how much is realistically attainable given that almost nowhere else on the west coast of North America is reporting SARs that achieve those levels. As we showed in our 2020 paper and expanded on in our response to the FPC in this document, coastwide SAR levels have largely fallen to numerically comparable levels to the Snake River.

Conclusions

The authors statement in results section 3.1 '*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 (Figure 2).*', is the basis for the major conclusions. They claim '*The policy implications of Chinook salmon SARs falling to about 1/3rd of early levels and converging to similar levels nearly everywhere along the west coast of North America are profound.*' They draw from this statement to make the following conclusions:

1. fisheries community need to re-assess several core conservation assumptions;
2. if survival also falls by roughly the same amount in regions with nearly pristine freshwater habitats (SE Alaska, north-central British Columbia), it is difficult to argue for a major role of regional factors in causing the decline;

3. question the actual effectiveness of freshwater habitat restoration initiatives when northern populations with nearly pristine freshwater conditions have similar SARs;
4. Given the similarity of the decline in survival, the economics of hatchery Chinook production are likely similar in other regions;
5. It is unclear whether the quality of reported harvest rate estimates is good enough for past PIT-based SAR estimates to be reliably converted into useful survival estimates;

All of these speculative conclusions drawn by the authors are rooted in fundamental problems in the basic CWT data they relied upon, the poor documentation of the source of these data, and the inconsistent application of these data in their estimates and surveys for SAR estimates. The authors do not document how they generated consistent estimates for SARs within and across regions. In fact, the authors conclude that 'we encountered substantial challenges in fully understanding whether all components of adult returns were adequately included in many SAR time series. This appears to be a fundamental flaw in the coastwide CWT based SAR analysis. This problem only compounds when the authors ignore the uncertainty for these CWT based SARs, then they attempt to evaluate patterns of survival across regions, and then compare and make judgement about the efficacy of much more precise PIT based SARs. The authors make no attempt at formally evaluating the similarity or difference in patterns of SAR estimates across populations. In fact, no attempt is made to estimate the CIs for any of the SAR estimates. The authors simply draw the conclusions for their interpretation of a visual inspection of figure 2. The conclusions drawn from similarities and differences from figure 3 do not provide an evaluation of the pattern of SARs. Also, given the differences in life histories, hatchery vs wild, and within hatchery groups experimental versus production groups the SAR comparisons are dubious at best.

KRS: All the SARs data in the paper are summarized and sourced in Table S1. The full dataset is available for download from the Dryad repository.

The problem with documenting the majority of the SAR estimates is that they are unpublished. We specifically used the PSC's Chinook Technical Committee indicator stocks because they are likely the best available. We supplemented this dataset with other estimates only after verification from hatchery biologists that there was no known reason why the estimates were not representative. Error on the CWT SAR estimates is likely impossible to calculate, but this should not prevent us from using these data. It would be useful to better evaluate the methods used to calculate SARs, which is why we explicitly called for the funding agencies to do so: "*We call for a systematic review by funding agencies to assess consistency and comparability of the SAR data generated and to further assess the implications of survival falling to similar levels in most regions of the west coast*".

More broadly, despite the issues with fully documenting the coastwide survival monitoring programs, all regions have fallen to about the same level. Schaller et al. have pointed out that the monitoring programs are complex and not well documented (which we agree with), but they have no explanation for why the reported SARs would decrease to about the same level.

The authors conclusion that ocean conditions drive these coastwide SAR patterns and freshwater conditions may not be drivers of SARs, is questionable given all the methodological flaws described above. However, the authors dubiously supported conclusions are dismissive of the large body of published literature that conclude Chinook salmon population survival rates are most influenced during estuary and early ocean life stages (Petrosky and Schaller 2010). Recruitment success in the ocean environment is generally believed to occur largely during the first critical months at sea (Ricker 1976; Nickelson 1986; Pearcy 1992; Mueter et al. 2002, 2005; Pyper et al.2005; Peterson et al. 2006). This early marine stage is a very vulnerable life stage for salmon because they make the transition from freshwater phase to a seawater phase involving numerous physiological changes while encountering marine predators. Therefore, especially recognizing the various dispersed ocean entry points for the salmon populations included in the Welch et al. paper, the influence of estuary and near shore ocean conditions would be highly variable (literature e.g., Peterman et al. shows strongest correlation of marine survival within about 500 km of ocean entry points), and would not support a similar influence across such a wide geographic range.

KRS: We have already discussed these issues in this response. Succinctly put, (1) "correlation is not proof of causation," and (2) assuming that the early life history phase is the critical driver of recruitment is an unproven assumption. We need to move to higher scientific standards of evidence rather than to rely on beliefs.

References

Haeseker, S.L. 2013. Nonrepresentative fish and ocean migration assumptions confound inferences in Rechisky et al. Proceedings National Academy of Science 110(37). www.pnas.org/cgi/doi/10.1073/pnas.1309087110

Haeseker, S.L., McCann, J.A., Tuomikoski, J., and Chockley, B. 2012. Assessing freshwater and marine environmental influences on life-stage-specific survival rates of Snake River spring–summer Chinook salmon and steelhead. Trans. Am. Fish. Soc. 141: 121–138. doi:10.1080/00028487.2011.652009.

ICTRT (Interior Columbia Technical Recovery Team). 2007. Viability criteria for application to Interior Columbia Basin Salmonid ESUs. ICTRT. Available: https://www.nwfsc.noaa.gov/research/divisions/cb/genetics/trt/col/trt_viability.cfm. (April 2020)

Mantua, N.J., Hare, S.R., Zhang, Y., Wallace, J.M. & Francis, R.C. 1997. A Pacific decadal climate oscillation with impacts on salmon. Bulletin of the American Meteorological Society 78: 1069–1079.
Marmorek, D. R., C. N. Peters, and I. Parnell, editors. 1998. PATH final report for fiscal year 1998. Prepared by ESSA Technologies, Vancouver.

McCann, J., B. Chockley, E. Cooper, B. Hsu, G. Scheer, S. Haeseker, R. Lessard, T. Copeland, E. Tinus, A. Storch, and D. Rawding. 2019. Comparative survival study (CSS) of PIT-tagged spring/summer/fall Chinook, summer steelhead and Sockeye. Annual Report to the Bonneville Power Administration, Contract 19960200, Portland, Oregon.

McCann, J., B. Chockley, E. Cooper, B. Hsu, S. Haeseker, R. Lessard, C. Petrosky, T. Copeland, E. Tinus, A. Storch, and D. Rawding. 2018. Comparative survival study (CSS) of PIT-tagged spring/summer/fall Chinook, summer steelhead and Sockeye. Annual Report to the Bonneville Power Administration, Contract 19960200, Portland, Oregon.

McCann, J., B. Chockley, E. Cooper, B. Hsu, H. Schaller, S. Haeseker, R. Lessard, C. Petrosky, T. Copeland, E. Tinus, E. Van Dyke, A. Storch, and D. Rawding. 2017. Comparative survival study (CSS) of PIT-tagged spring/summer/fall Chinook, summer steelhead and Sockeye. Annual Report to the Bonneville Power Administration, Contract 19960200, Portland, Oregon

Mueter, F.J., Peterman, R.M. & Pyper, B.P. 2002. Opposite effects of ocean temperature on survival rates of 120 stocks of Pacific salmon (Oncorhynchus spp.) in northern and southern areas. Canadian Journal of Fisheries and Aquatic Sciences 59: 456–463.

Mueter, F.J., Pyper, B.J. & Peterman, R.M. 2005. Relationships between coastal ocean conditions and survival rates of Northeast Pacific salmon at multiple lags. Transactions of the American Fisheries Society 134: 105–119.

Nickelson, T.E. 1986. Influences of upwelling, ocean temperature, and smolt abundance on marine survival of coho salmon (Oncorhynchus kisutch) in the Oregon Production Area. Canadian Journal of Fisheries and Aquatic Sciences 43: 527–535.

Pearcy, W.G. 1992. Ocean ecology of North Pacific salmonids. Seattle, WA: University of Washington Press.

Peterman, R. M. B. J., M. F. Pyper, M. D. Adkison Lapointe, and C. J. Walters. 1998. Patterns of covariation in survival rates of British Columbian and Alaskan Sockeye Salmon (Oncorhynchus nerka) stocks. Canadian Journal of Fisheries and Aquatic Sciences 55:2503–2517.

Peterson, W.T., Hooff, R.C., Morgan, C.A., Hunter, K.L., Casillas, E. & Ferguson, J.W. 2006. Ocean conditions and salmon survival in the Northern California Current. Newport, OR: Northwest Fisheries Science Center, Newport Research Station. Available at: http://www.nwfsc.noaa.gov/research/ divisions/fed/ecosysrep.pdf.

Petrosky, C.E., and Schaller, H.A. 2010. Influence of river conditions during seaward migration and ocean conditions on survival rates of Snake River Chinook salmon and steelhead. Ecol. Freshw. Fish, 19(4): 520–536. doi:10.1111/ j.1600-0633.2010.00425.x

Petrosky, C.E., H.A. Schaller, E.S. Tinus, T. Copeland and A.J. 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. North American Journal of Fisheries Management 40:789–803.

Pyper, B.J., Mueter, F.J. & Peterman, R.M. 2005. Acrossspecies comparisons of spatial scales of environmental effects on survival rates of Northeast Pacific salmon. Transactions of the American Fisheries Society 134: 86–104.

Ricker, W.E. 1976. Review of the rate of growth and mortality of Pacific salmon in salt water, and noncatch mortality caused by fishing. Journal of the Fisheries Research Board of Canada 33: 1483–1524.

Schaller, H.A., and Petrosky, C.E. 2007. Assessing hydrosystem influence on delayed mortality of Snake River stream-type chinook salmon. N. Am. J. Fish. Manage. 27: 810–824. doi:10.1577/M06-083.1.

Schaller, H., P. Wilson, S. Haeseker, C. Petrosky, E. Tinus, T. Dalton, R. Woodin, E. Weber, N. Bouwes, T. Berggren, J. McCann, S. Rassk, H. Franzoni, and P. McHugh. 2007. Comparative survival study (CSS) of PIT-tagged spring/summer Chinook and steelhead in the Columbia River basin: ten-year retrospective summary report. Report to the Bonneville Power Administration, Projects 1996-020-00 and 1994-033-00, Portland, Oregon.

Schaller, H. A., C. E. Petrosky, and E. S. Tinus. 2014. Evaluating river management during seaward migration to recover Columbia River stream-type Chinook Salmon considering the variation in marine conditions. Canadian Journal of Fisheries and Aquatic Sciences 71:259– 271.

Kintama Citations

Hjort, J. (1914). Fluctuations in the great fisheries of northern Europe viewed in the light of biological research. *Rapp. P.-V. Reun. Cons. Int. Explor. Mer, 20*, 1–228.

Horton, R. (2015). Offline: What is medicine's 5 sigma. *Lancet, 385*(9976), 1380. doi:10.1016/S0140-6736(15)60696-1

Rechisky, E. L., Welch, D. W., & Porter, A. D. (2013). Reply to Haeseker: Value of controlled scientific experiments to resolve critical uncertainties regarding Snake River salmon survival. *Proc. Nat. Acad. Sci. USA*. doi:10.1073/pnas.1310239110

Rechisky, E. L., Welch, D. W., Porter, A. D., Hess, J. E., & Narum, S. R. (2014). Testing for delayed mortality effects in the early marine life history of Columbia River yearling Chinook salmon. *Mar. Ecol. Prog. Series, 496*, 159–180 doi:10.3354/meps10692